

NOĒSIS

UNDERGRADUATE JOURNAL OF PHILOSOPHY
AT THE UNIVERSITY OF TORONTO

XVI

**Pinocchio Craps and Robin Hood Hold'em: Betting
on the First Formula of the Categorical Imperative**
Arash Ghiassi

**Existence versus Primacy: A Critical
Overview of Phenomenal Intentionality**
Michael Galang & Can Mekik

**The Telltale Footprint: Doctored Evidence and
Epistemic Value**
Ryan Martins

**Hinging Truth: An Examination of Truth and Falsity of
Wittgensteinian 'Hinge-Propositions'**
Christopher Sullivan

A Defense of Abortion Rights on Lockean Principle
Edward Ruse

Classics and Philosophy
In conversation with Brad Inwood

Philosophy and Obligation
In conversation with Ashley E. Taylor

NOĒSIS

(Greek: intelligence, thought, understanding, mind). In Greek philosophy, the knowledge that results from the operations of *nous* (the mind, reason, intellectual faculty).

NOĒSIS
UNIVERSITY OF TORONTO
UNDERGRADUATE JOURNAL OF PHILOSOPHY

XVI

Editors-in-Chief
C. Dalrymple-Fraser
Rory Harder

Managing Editor
SuJung Lee

Senior Editor
T. R. Abrahams

Associate Editors
Clara Christopher
Robert Fraser
Joy Shim

Contributors
Michael Galang
Arash Ghiassi
Ryan Martins
Can Mekik
Edward Ruse
Christopher Sullivan

Spring 2015

NOĒSIS
UNDERGRADUATE JOURNAL OF PHILOSOPHY
AT THE UNIVERSITY OF TORONTO

XVI

Papers

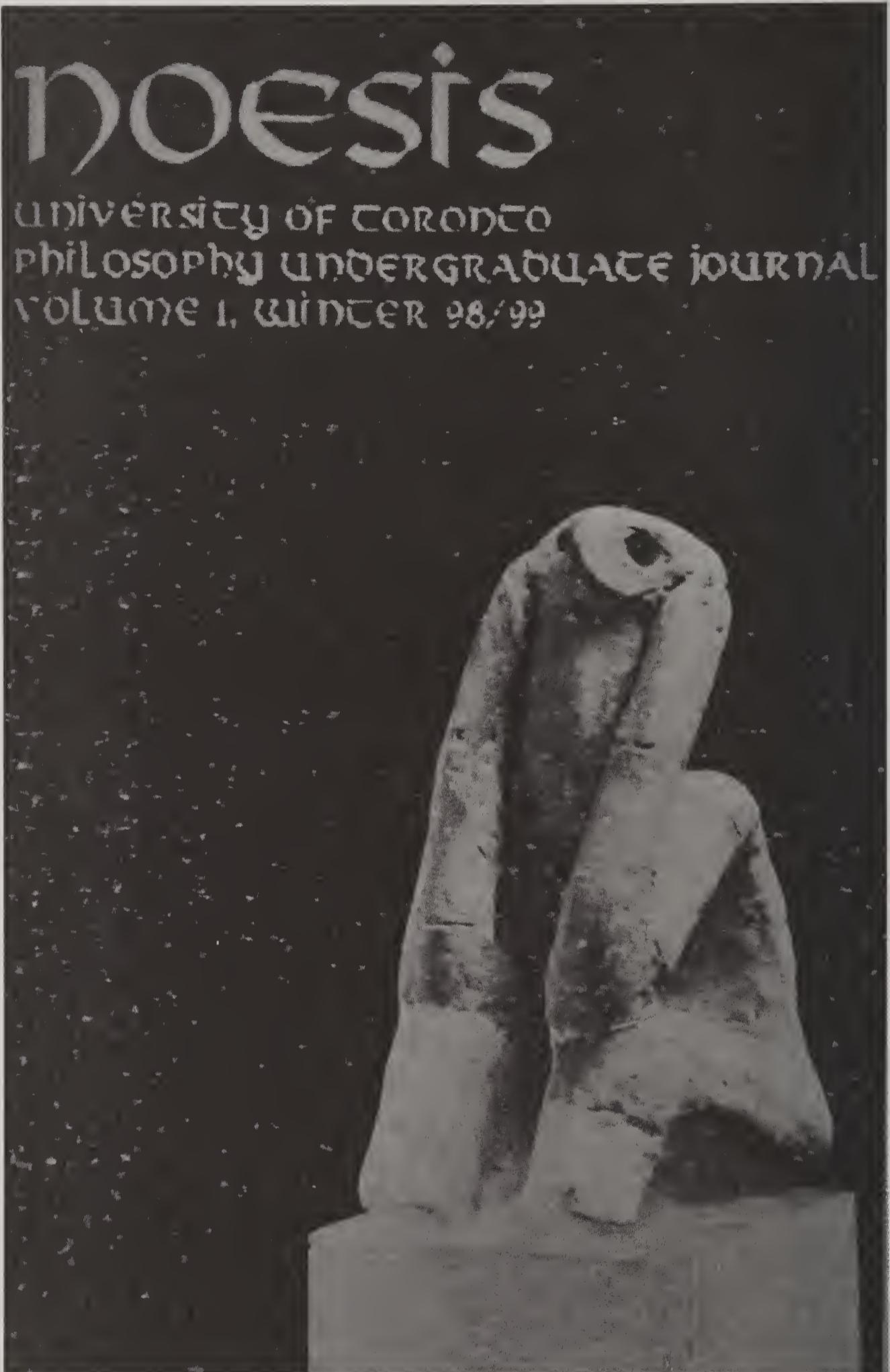
Pinocchio Craps and Robin Hood Hold’em: Betting on the First Formula of the Categorical Imperative.....	7
Arash Ghiassi	
Existence versus Primacy: A Critical Overview of Phenomenal Intentionality	23
Michael Galang and Can Mekik	
The Telltale Footprint: Doctored Evidence and Epistemic Value	41
Ryan Martins	
Hinging Truth: An Examination of the Truth and Falsity of Wittgensteinian 'Hinge-Propositions'	55
Christopher Sullivan	
A Defense of Abortion Rights on Lockean Principles.....	67
Edward Ruse	

Interviews

Classics and Philosophy	77
Rory Harder and SuJung Lee in conversation with Brad Inwood	
Philosophy and Obligation	89
Charles Dalrymple-Fraser in conversation with Ashley E. Taylor	

noēsis

UNIVERSITY OF TORONTO
PHILOSOPHY UNDERGRADUATE JOURNAL
VOLUME I, WINTER 98/99



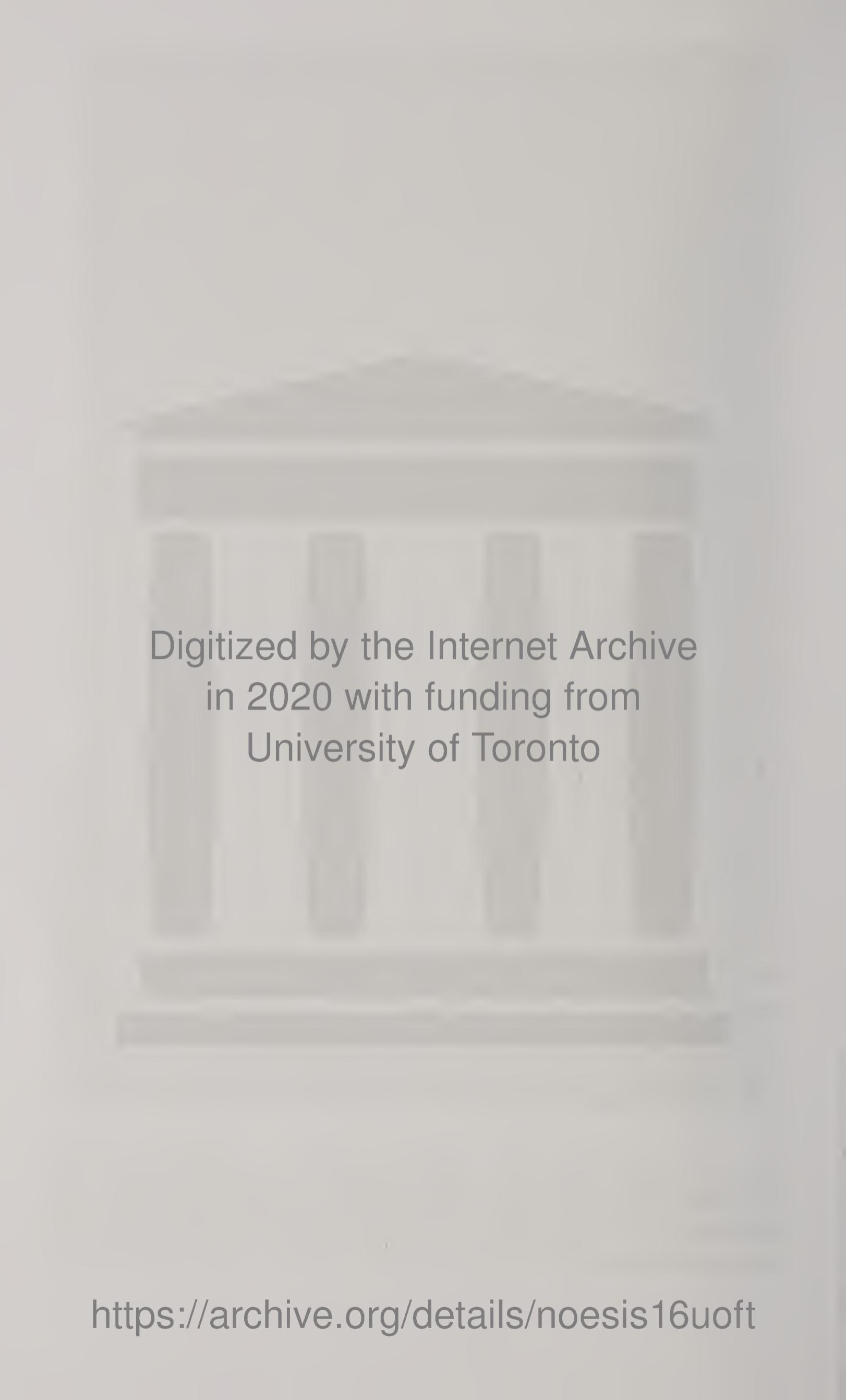
Above: The cover of the first issue of Noēsis. The cover, designed by Ramona Ilea, features the famous Cucuteni-Trypillian statue “The Thinker of Tarpești” (Gânditorul din Târpești). The statue dates around 4,500BCE, and is our oldest known “thinker”. The Thinker of Tarpești adorned the covers of Noēsis I–V.

Editor's Note

Welcome to *Noēsis*, the University of Toronto's undergraduate journal of philosophy. The aim of the journal is to provide a forum for philosophical work and discourse, and to foster a culture of lively philosophical exchange within the University of Toronto's undergraduate community. *Noēsis* is an undergraduate publication: from the original articles to the peer review process, editing and production, every aspect of the journal is the result of work done by our own undergraduates.

There are many individuals to whom *Noēsis* is greatly indebted for their continued support, and without whom the journal would likely cease to exist. We'd like to especially thank the officers of the Philosophy Department of the University for their continuing support of the journal and guidance for its editors. We also owe our gratitude to U of T philosophy faculty at large for encouraging talented students to submit papers for publication, and to those many students who submitted to this issue of the journal. Finally, we owe our thanks to those professors who generously shared their time and thoughts with us for our interviews in this issue: Professors Brad Inwood and Ashley Taylor.

This issue of *Noēsis* was made possible by generous financial assistance from the Arts and Science Student's Union, the Department of Philosophy, St. Michael's College, and University College. Thank you.



Digitized by the Internet Archive
in 2020 with funding from
University of Toronto

<https://archive.org/details/noesis16uoft>

Pinocchio Craps and Robin Hood Hold'em: Betting on the First Formula of the Categorical Imperative

Arash Ghiassi

Interpreting the relation between Kant's three formulae of the Categorical Imperative has been subject of debate between scholars, but the first of these formulae—i.e. the Formula of the Universal Law—is widely held in a privileged position. It is often thought to be the best-known formula, the main one, or the one most suitable for precise action-guiding.¹ Yet, providing a satisfactory interpretation of the First Formula itself is also a site of disagreement. While some contemporary Kant scholars have found a *practical* interpretation of this formula appealing,² in this essay I will cast doubt on whether even it will lead to satisfactory consequences.

Describing the Formula of the Universal Law, Kant says, “there is therefore only a single categorical imperative, and it is this: *act only according to that maxim through which you can at the same time will that it become a universal law*” (G4:421,³ emphasis in the original). This gives rise to a handy test for someone requiring moral guidance: she merely has to see whether or not there is a conflict between acting in accordance with her maxim on the one hand and at the same time willing

¹ Onora O’Neil, *Acting on Principle: An Essay on Kantian Ethics* (New York: Columbia University Press, 1975), 1; 32. See also Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), 106.

² Such as Korsgaard, as I will show in this paper.

³ Immanuel Kant, *Groundwork of the Metaphysics of Morals*, ed. Mary Gregor and Jens Timmermann (Cambridge: Cambridge University Press, 2012). For ease of reference, I will cite the *Groundwork* in this fashion throughout the paper.

that it become a universal law on the other. If there is, she may not proceed; if there is not, she may. This is intuitive: if she acts in a way that she cannot at the same time will others to act, she is in effect unreasonably viewing herself as an exception to a rule of conduct she has to acknowledge for everyone.

But if the Formula of the Universal Law is going to be an acceptable action-guiding principle, this test must lead us in the right direction every time. A charitable interpretation of this proposed test, then, is crucial in evaluating its action-guiding merits. Interpreted uncharitably, the test might unfairly appear to be an unreliable moral touchstone. In Part (I), I will survey some of the ways in which the test has been interpreted. In Part (II), I will present some examples of situations which, I think, would undermine even the most charitable interpretation found in (I). In Part (III), I will deal with a potential problem for my examples that could arise based on the problem of relevant descriptions. Finally, I will offer some concluding remarks in (IV).

I. Various Interpretations of the First Formula

Kant provides two kinds of possible conflicts that may arise upon applying the Formula of the Universal Law to an impermissible action:

Some actions are such that their maxim cannot even be thought without contradiction as a universal law of nature; let alone that one could will that it should become such. In the case of others that inner impossibility is indeed not to be found, but it is still impossible to will that their maxim be elevated to the universality of a law of nature, because such a will would contradict itself (G4:424).

In other words, the application of the test to the former group of impermissible maxims leads to a *contradiction in conception* and corresponds to unrelenting duties (of omission), while its application to the latter involve a *contradiction in will* and corresponds to meritorious

duties (G4:424). In this paper I am mainly concerned with unrelenting duties and interpretations of the way in which a *contradiction in conception* is supposed to arise. Clarifying what is meant by a contradiction in conception is the aim of this section.

It is useful to ground our investigation in the second case study provided by Kant in which an unrelenting duty arises. In this example a financially broke moral agent finds himself facing the possibility of adopting the following maxim: "... when I believe myself to be in need of money I shall borrow money, and promise to repay it, even though I know that it will never happen" (G4:422). According to Kant, this person—let us call him Pinocchio for convenience—ought not to adopt such a maxim, because he cannot behave in this way while at the same time willing that others do so as well. Kant explains that Pinocchio's maxim, if universalized and imposed on everyone as a rule, would make promises impossible. Kant does not provide a precise definition of 'promise' beyond the common-sense understanding of the term. He states, however, that if everyone made a dishonest representation in order to get their hands on some cash in times of need, no one would take that sort of representation seriously as a promise (G4:422). Pinocchio's maxim to *make a kind of promise*, then, presupposes that this is not the case, and therefore he cannot will both to make a lying promise—that is, a promise that is simultaneously a lie—and that everyone else make them too. That would lead to a certain kind of contradiction.

But what is the precise sense in which Pinocchio's maxim already presupposes such a rule of conduct, and how does the contradiction arise? There are at least three ways, broadly speaking, in which this contradiction is interpreted in the Kantian literature. These are: the teleological interpretation, the logical interpretation, and the practical interpretation.⁴ I will present each of the interpretations here, but as I will argue in later sections, I do not think that the Formula of the Universal Law provides an acceptable action-guiding test under any of them.

Let us first take a look at the teleological interpretation. According to this view, the test in the First Formula is to see whether the proposed maxim would defeat its own purpose if universalized. O'Neil's

⁴ Korsgaard, *Kingdom of Ends*, 78.

view, for instance, seems to be a version of this view. According to her, if one intends to do anything, one has to also intend “some sufficient set of conditions to realize [one’s] ends and the *normal, predictable* results” of that act.⁵ In assessing the preconditions and consequences of an action, we must assume that we belong to the same system of nature to which people actually belong—where causal relations are such as we know them.⁶ Pinocchio’s lying promise has a purpose: inducing the moneylender to give him a loan. If Pinocchio intends to make such a lying promise to that purpose, he must also intend that (i) preconditions for promising, including a climate of trust, be in place. On the other hand, if he intends his maxim as a universal law, he also intends that (ii) the predictable consequences of everyone making lying promises, including the erosion of a climate of trust, come to pass. But (ii) undermines (i), and therefore there is a contradiction in willing both. The maxim to make a false promise, then, fails the test because promising fails to bring about its intended purpose when universalized.⁷

Paton provides a different account of the teleological interpretation. According to him, we must consider all phenomena as if they have natural ends or purposes.⁸ In particular, we should understand human nature and actions as structured according to a systematic harmony of ends. For the limited purposes of this section, let us grant Paton that a plausible account of this kind of natural harmony is available, without getting into details. Accordingly, we must “ask whether any proposed maxim, if it were made a law of nature, would fit into such a systematic harmony.”⁹ Only if the maxim would fit in this harmony is it permissible to act on. In the case of Pinocchio, according to this view, we should recognize that a promise *qua promise* has a natural purpose—namely that of promoting trust and cooperation. Given current causal laws, however, this purpose would no longer be attainable by

⁵ O’Neil, *Acting on Principle*, 70. Emphasis added.

⁶ *Ibid.*, 74.

⁷ *Ibid.*, 78.

⁸ Herbert J. Paton, *The Categorical Imperative: A Study in Kant’s Moral Philosophy* (London: Hutchinson, 1967), 149.

⁹ *Ibid.*, 150.

promises if making lying promises were a law of nature. As a result, any law of nature to that effect would be contradictory insofar as it would defeat a natural purpose inherent in it.¹⁰ This would not be compatible with a systematic harmony of ends, and therefore Pinocchio's maxim is an impermissible one.

But Kant thinks that Categorical Imperatives are synthetic *a priori* (G4:420) and therefore *necessarily true*. The teleological interpretation, however, would make the moral status of false promises contingent on particular causal relations that may or may not obtain. According to that interpretation, in a possible world in which a lying promise has different consequences or preconditions than those mentioned above, making a lying promise may not be impermissible. So, the claim that promises will eventually stop being convincing as a causal consequence of widespread lies does not explain why lying promises are impermissible for all possible rational agents. For instance, in a possible world populated by gullible moneylenders whose trust can be won with a smile and a warm cup of coffee, the general climate of trust may never fully erode as a causal consequence of a widespread practice of lying promises. Promise-making may therefore never die and its purpose may never be defeated. Yet, lying promises would still be wrong.

This difficulty motivates the logical interpretation, to which I now turn. According to this interpretation, which is endorsed by Kemp, in no possible world does the proposed maxim prevail as law. Such an impossible world, according to Kemp, would include a logical contradiction, namely that there both be and not be promises.¹¹ On the one hand, the universalized law that everyone can make a lying promise in order to enrich themselves presupposes that there are such things as promises to be made. On the other, prevalence of lying promises would entail that promises do not exist at all, since the practice of promise-making presupposes an environment of trust. This entailment, Kemp stresses, “must be one of logical, or quasi-logical, not causal consequence.”¹²

¹⁰ *Ibid.*, 153.

¹¹ J. Kemp, “Kant's Examples of the Categorical Imperative”, *The Philosophical Quarterly*, Vol. 8, No. 30 (Jan 1958), 67.

¹² *Ibid.*

But we may remain skeptical, as O’Neal does, as to whether such a purely logical contradiction arises in the case of promise-making.¹³ But again, let us grant Kemp that some plausible account of a logical contradiction is in principle available. Nonetheless, any such account of a logical interpretation of the Formula of Universal Law faces the following disadvantage: it marks some seemingly permissible maxims as contrary to duty.¹⁴ For instance, the maxim to buy carrots but not sell them seems perfectly innocuous. Yet, its universalization would lead to a logical absurdity. In no possible worlds can agents universally follow the rule “buy, but don’t sell, carrots” since every purchase requires a sale—and that is an analytic, and not a merely contingent, truth. Yet, despite this logical contradiction, it hardly seems to be an unrelenting duty not to act according to that maxim.

Lastly, the practical interpretation avoids this and other problems associated with the logical and the teleological interpretations. According to Korsgaard’s presentation of this interpretation, the world in which the agent’s impermissible maxim reigns as law is not impossible, but is one in which she cannot act on the maxim by taking the specified means in order to achieve the specified end.¹⁵ Therefore, so long as she wills her maxim, she cannot also will that it be a universal law, because that would defeat the purpose inherent in her maxim.¹⁶ Yet, the way in which her purpose would be defeated is not contingent on causal laws in this interpretation. Rather, the universal adoption of the means specified in the maxim would analytically entail that it is no longer available as a means to the specified end.

In no possible world in which agents universally make lying promises, for instance, can promises be the means of achieving anything at all, according to this interpretation. This is because promise-making is possible only as an institution within which words normally carry weight, unlike what the universal law to make lying promises entails.¹⁷

¹³ O’Neil, *Acting on Principle*, 66.

¹⁴ *Ibid.*, 68.

¹⁵ Korsgaard, *Kingdom of Ends*, 92.

¹⁶ *Ibid.*, 93.

¹⁷ *Ibid.*, 64.

Promises just are takings-on of commitments, and therefore a law to the effect that promises universally fail to reflect commitments necessarily negates the institution of promise-making. The universalization of the maxim to make a lying promise, in Kant's words, "would make the promise and the end one may pursue with it itself impossible, as no one would believe he was being promised anything, but would laugh about any such utterance, as a vain pretence" (G4:422).

Kant does not say that no one would believe that the promise he was made was sincere, but that *he was promised anything at all*. One cannot make what could properly be called a 'promise' under the proposed universal law, even if one utters all the words and goes through all the motions. What one performs is necessarily something else. Therefore, precisely because Pinocchio intends to take up as his means a sort of promise—and not, say, a meaningless song-and-dance, which is perhaps morally permissible—he cannot consistently will that lying promises be the norm. Or at least, that is the practical interpretation of the application of the First Formula to Pinocchio's maxim.¹⁸

II. Pinocchio Craps and Robin Hood Hold'em

Consider the following fictitious game of *Pinocchio Craps*. Here is how it works: each person can only play once in her lifetime, presumably when faced with some grave financial difficulty. The player is presented with n dice, where $n > \log_6(N)$ and $N =$ the current population of the Earth.¹⁹ (Today this means that there have to be at least 13 dice). Each die has only one side marked 'Yes', every other side being marked 'No'. Since each die has six sides, the probability of throwing n 'Yes's is at most $\frac{1}{N}$. Still, if and only if the player gets that outcome, she will proceed to make a false promise to a non-suspecting moneylender in order to borrow money she does not intend to pay back. Given the number of dice involved, chances are that at most one person in the entire Earth will

¹⁸ *Ibid.*

¹⁹ This choice of N is largely heuristic, as it readily provides a probability which is intuitively suitable for my purposes. The same point could be made with "N = the number of rational occupants of the relevant spatiotemporal window" without begging any questions about who the relevant window actually includes. But that would be too abstract a way of introducing my counter-example.

actually make a lying promise—even if everyone played. Is playing Pinocchio Craps in order to alleviate financial pressure permissible?

The correct answer seems to be no. A lying promise is intuitively wrong in virtue of something which is still present even if the promise is made with a lower probability. Moreover, playing Pinocchio Craps is clearly in violation of Kant's Second Formula of the Categorical Imperative: "*So act that you use humanity, in your own person as well as in the person of any other, always at the same time as an end, never merely as a means*" (G4:429). Like the simple lying promise case mentioned by Kant, during the course of this game the moneylender is intended to be used as a mere means. This is because information crucial to his free consent about the transaction is concealed from him and so he cannot at the same time act as an end of the transaction by agreeing to it (4:430).

The First Formula, however, seems to permit playing Pinocchio Craps under all available interpretations. Here one can will both to play the game and at the same time that playing the game be a universal law. There is no teleological inconsistency, because everyone else's involvement in the game does not alter one's chances of getting a flush 'Yes'. And, since almost no one else actually gets it, it will not causally affect the believability of one potential false promise. Indeed, promises are still generally believed despite the fact that many insincere ones are also made at present. Hence, the efficacy of the game or of promise-making to bring about the relevant ends will not be diminished. There is no logical inconsistency, because a world in which everyone plays Pinocchio Craps seems perfectly conceivable, as an extremely low probability of a lying promise does not entail that there both be and not be promises. Similarly, there is no practical inconsistency, because a universal law permitting only very unlikely deviations from honesty does not mean that promise-making as a whole is nonexistent. Since the Pinocchio Craps player need not will that promises are generally made insincerely, she can at the same time will to make a lying promise as a means to her end in the event that she gets lucky.

Here is another example: a reincarnation of Robin Hood wants to steal money from those richer than him in order to promote equality,

while at the same time holding on to what he steals. This modus operandi—let us call it *Robin Hood Hold 'em*—is consistent: if it is successfully carried out, Robin Hood would be as rich as the other richest person. But this outcome would be more equal than the status quo, because the richest person would be less rich and the amount of wealth would be shared with more people at the top. Is the maxim permissible?

Again, the correct answer seems to be no, both intuitively and according to the Second Formula. The people whose money is taken by Robin Hood do not consent to the transaction and are thus merely used as means to his end. Robin Hood, therefore, is a “transgressor of the rights of human beings” and his actions are “attacks on the freedom and property of others” (G4:430). But does the First Formula of the Categorical Imperative also condemn his actions?

And again, it seems that it does not. In a world where everyone steals from only those who are richer, everyone will eventually have an exactly equal amount of money. It seems, then, that universalizing Robin's maxim actually makes it more likely that his end—equality—be brought about, not less likely. Therefore, there seems to be no teleological nor logical contradiction. In this regard, it is like helping the poor: if the maxim to help the poor was universalized, there would be no more poor people to be helped. But this does not mean that the maxim has been defeated, rather that it has succeeded.²⁰ It is not logically contradictory to help all poor people in a special case where there are not any, in the same way that it is consistent to steal from any richer person in a world without any. There is also no practical contradiction, because widespread stealing from those richer than oneself does not mean that stealing will be unavailable as a means to be taken up by Robin. I will explain this last point more fully in following paragraphs.

There is a standard Kantian objection to any maxim involving theft which is ostensibly open to all three interpretations of the supposed contradiction in conception. It is often argued that a thief wills that a system of property both exists and does not exist if he is to will that his

²⁰ Korsgaard, *Kingdom of Ends*, 87.

maxim at the same time be a universal law.²¹ Insofar as he intends to own the things he takes, he wills that a system of property ownership be in place. Insofar as he intends to take them away from their previous owner, he wills to undermine and destroy that same system of property. Therefore, it is argued, the First Formula rules against indiscriminate theft and looting as a self-contradictory maxim. This contradiction, as is expected, could be cast in a logical, teleological, or practical light. I will not get into the details of each interpretation of this particular contradiction, because I think that no contradiction can arise in the case of Robin Hood Hold 'em at all.

The difference here is that the thefts are not indiscriminate. They only occur in certain circumstances—namely, where the previous owner of the stolen money is richer than the subsequent one. Therefore I think that Robin Hood does not simultaneously will the existence and non-existence of the same system of property. Rather, he wills the non-existence of the usual system of property and the existence of an alternative system; call this alternative *asymmetric property*.

In this alternative system everyone has the conventional right to keep their asymmetric property and try to prevent others from taking it. What is more, no one has a right to take the asymmetric property of others who are as poor as or poorer than they are. However, everyone has a right to try and take the asymmetric property of those with more of it than they have. So whereas in the usual system of property everyone has immunity against everyone else taking their property, here you only have such immunity against those equally as rich as or richer than you. If those poorer than you manage to take your asymmetric property it becomes rightfully theirs, in exactly the same asymmetrical sense as what you had was rightfully yours. But you can try to prevent that, for example by investing in heightened security. This is not necessarily an attractive system of property, but all that matters is that it is possible and Robin Hood can will it. Therefore, no contradiction arises concerning his maxim. In particular, there is no practical contradiction, because even if

²¹ Marcus G. Singer, *Generalization in Ethics: An Essay in the Logic of Ethics, with the Rudiments of a System of Moral Philosophy* (New York: Knopf, 1961), 252. See also Korsgaard, *Kingdom of Ends*, 98.

everyone followed it, Robin Hood would still be able to take up the specified means—i.e. stealing—in order to achieve his end.

III. The Problem of Gerrymandering

There is a general worry within deontological ethics which might surface as a problem for my counter-examples. In this section I will first explain this worry before exploring its possible implications for my position. The worry is that it is not always clear which description of an action is the one to be used in its moral evaluation. The same action could be described as speaking, lying, lying to save a life, lying to save a villain's life, lying to save a villain's life on Sunday, etc. Each description, moreover, might turn out to have different duties attached to it. As O'Neil puts it, "Any instantiation of a principle is an instantiation of many principles, and any principle has many instantiations."²² As such, she thinks that any universality test can be action-guiding only if it passes the same judgment of permissibility about an action under all of its relevant descriptions.²³ But which descriptions are relevant?

Certainly, descriptions whose scopes are artificially restricted cannot be relevant, because otherwise any action would turn out to be permissible. For instance, we might imagine Pinocchio arguing, "Seeing as no one else is around, my maxim to make a lying promise in exactly ten seconds while standing in this particular corner of the universe is perfectly permissible to act on, because even if it was universalized false promise-making would remain possible." Here, the maxim is specified in such a way that it arbitrarily limits the moral agents to whom the universalized law would apply, and passes the test only for that reason. These kinds of descriptions provide a way of cheating around the First Formula by "gerrymandering" a portion of the Kingdom of Ends to one's advantage. I have not yet provided an exact analysis of gerrymandering, but we must exclude all gerrymandered descriptions if we are to have a coherent test.

Accordingly, in order to anticipate the objection that my counterexamples must also be excluded, I need to show that they are not unduly gerrymandered. But in order to exclude gerrymandered

²² O'Neil, *Acting on Principle*, 11.

²³ *Ibid.*, 12.

descriptions, we must first find a general definition of gerrymandering that provides a way of identifying its instances according to a criterion that establishes which descriptions of an action are relevant.

O’Neil, for example, thinks that Kant provides just the needed account of such a criterion: the only relevant description of an agent’s action is the one she chooses to act on—namely, her maxim.²⁴ Every voluntary act, then, has a maxim that can be evaluated using the universalizability test. In order to block the possibility of gerrymandering O’Neil appeals to honesty in representing one’s true maxim. “A person cannot simply *claim* a highly specific maxim,” she explains, “he must, in fact, intend his act to be contingent on those restrictions.”²⁵ This means that if Pinocchio’s claim that his maxim is one of “making a lying promise in order to receive money in exactly ten seconds while standing in this particular corner of the universe” is to be taken seriously, he must genuinely intend to act according to the precise restrictions specified in it. If he would make the lying promise despite a change in the spatiotemporal context, it means that he was not honestly taking up that maxim and was instead using it as a loophole for circumventing the moral law. Gerrymandering, then, is just a description of one’s action that is marked by this kind of dishonesty, according to O’Neil. Her appeal to honesty, however, is not made because of any moral value attached to truthfulness, but because it helps us uncover and judge what the maxim really was. Knowing Pinocchio, we know that his lies are really contingent on no such spatiotemporal restrictions, and therefore the claimed restrictions on his maxim should be ignored.

However, the Robin Hood in my example does avoid stealing from the poor. And whoever plays Pinocchio Craps will really avoid making a false promise if she loses, because otherwise she would not have gone through the trouble of playing it in the first place. Perhaps she is actually going through that trouble precisely in order to conform to a Kantian moral constraint. It seems, therefore, that the added descriptions in their stated maxims could both be an honest representation of their

²⁴ *Ibid.*, 13.

²⁵ *Ibid.*, 72. Emphasis in original.

intentions. This account of the criterion of relevance therefore seems not to implicate these examples in gerrymandering.

I will now look at another such account provided by Singer, who also acknowledges the problem of gerrymandering.²⁶ He notes that all these cases involve arbitrary references to particulars, relative to which the description can be *reiterated*.²⁷ For instance, Pinocchio can make his gerrymandered argument over and over with respect to any time and place whatsoever. This signifies that his argument does not depend on the specific information provided, but only that *some particular information be provided*. By contrast, if reference to a particular is really relevant to the universality test, the description containing that reference would not be “reiterable” with respect to it. Reiterable descriptions, Singer concludes, must be rejected as gerrymandered.²⁸

That Robin Hood Hold'em does not include reiterable references to particulars is easy to see: it is stated in very general terms, and with no reference to a particular time, place, or victim. The case is trickier when it comes to Pinocchio Craps, because it does include a reference to a particular turn of events, namely a rare occurrence of a flush ‘Yes’ as opposed to other possible outcomes of the dice throw. However, the argument for permissibility of Pinocchio Craps is not reiterable with other possible combination of the dice. This is because the probability of other such combinations is higher than the one connected with a flush ‘Yes’, while the argument hinges on the extremely low probability of a false promise actually being made as part of the game. It seems, therefore, that this instance of a reference to a particular is actually relevant to the universality test and cannot be ignored as spurious. Singer’s criterion, too, seems to allow both of the examples in (II).

IV. Conclusion

If I am right about Pinocchio Craps and Robin Hood Hold'em, then that seems to tell against an interpretation of the Categorical Imperative such as Skorupski’s, according to which it is intended to yield “*all and only*

²⁶ Singer, *Generalization in Ethics*, 140.

²⁷ *Ibid.*, 82.

²⁸ *Ibid.*, 83.

the universal principles of morality.”²⁹ More accurately, it tells against such an interpretation only if we also think that the First Formula is an adequate expression of the Categorical Imperative—as Skorupski does while additionally taking all the formulas to be intended as equivalent.³⁰ As it stands, the First Formula seems to yield some false permissions as principles of morality, in clear disagreement with the second formula. It seems, in other words, that First Formula is not the more precise action-guider of the three, in contrast with O’Neil and Korsgaard’s interpretations.

It is, however, more consistent with an alternative interpretation put forth by Allen Wood, who interprets the First Formula not as the main or the most precise one, but rather as the most provisional.³¹ Far from being equivalent, Wood allows that the different formulae are “objectively” different—i.e. command different duties³²—but also disagrees that these duties can be simply *deduced* from any of the formulae. Rather, he thinks that these duties may only be found through a much looser process of interpretation in light of other considerations and based on an Aristotelian *phronesis*.³³ Taking this approach, however, would diminish the informative nature of Kantian thought with regards to moral duties.

²⁹ John Skorupski. “Autonomy and Impartiality: Groundwork III,” in *Kant’s Groundwork of the Metaphysics of Morals: A Critical Guide*, ed. Jens Timmermann (Cambridge, UK: Cambridge University Press, 2009), 159. Emphasis in original.

³⁰ *Ibid.*

³¹ Allen W. Wood, *Kantian Ethics* (Cambridge: Cambridge University Press, 2008), 66.

³² *Ibid.*, 69.

³³ *Ibid.*, 60; 64.

Works Cited

Kant, Immanuel. *Groundwork of the Metaphysics of Morals*. Edited by Mary Gregor and Jens Timmermann. Cambridge: Cambridge University Press, 2012.

Kemp, J. "Kant's Examples of the Categorical Imperative." *The Philosophical Quarterly*, Vol. 8, No. 30 (Jan., 1958): 63-71

Korsgaard, Christine M. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.

O'Neill, Onora. *Acting on Principle: An Essay on Kantian Ethics*. New York: Columbia University Press, 1975.

Paton, Herbert J. *The Categorical Imperative: A Study in Kant's Moral Philosophy*. London: Hutchinson, 1967.

Singer, Marcus G. *Generalization in Ethics: An Essay in the Logic of Ethics, with the Rudiments of a System of Moral Philosophy*. New York: Knopf, 1961.

Skorupski, John. "Autonomy and Impartiality: Groundwork III." In *Kant's Groundwork of the Metaphysics of Morals: A Critical Guide*. Edited by Jens Timmermann. Cambridge, UK: Cambridge University Press, 2009.

Wood, Allen W. *Kantian Ethics*. Cambridge: Cambridge University Press, 2008.

Existence Versus Primacy: A Critical Overview of Phenomenal Intentionality

Michael Galang & Can Mekik

Phenomenal intentionality challenges the dominant approach to intentionality in the analytic tradition. There are two theses that characterize phenomenal intentionality: that it exists, and that traditional theories of intentionality are not equipped to account for it; that is to say, that it is somehow basic.¹ We argue that the existence of phenomenal intentionality is plausible, but that the second thesis is suspect. Phenomenal intentionality presupposes that both phenomenology and intentionality are real psychological phenomena, and its defenders present it as a competitor with other realist approaches to the above. We therefore discuss both intentionality and phenomenality as realists would. Our discussion is in three major parts. We first make some general remarks about intentionality; we then discuss what Horgan and Tienson (2002) call separatism and their denial of it; and, finally, we put forward our argument for the thesis stated above.

I. A Brief Overview of Intentionality

The classical paradigms for intentionality are all rooted in the analysis of beliefs, desires, hopes, fears, and the like. If there is one theoretical feature that the classical paradigms share, it is that their point of departure is the construal of intentionality as a feature that warrants or enables a special kind of relation.

One theoretical tradition (Brentano, Twardowski, Meinong and followers) holds that every intentional state has an object (Sajama and

¹ By traditional ‘theories of intentionality,’ we mean the theories discussed in Fodor (1994). Also see ‘The Primacy Thesis’ below.

Kampinen, 1987). This claim is at the heart of the most famous puzzle of intentionality: there are intentional states whose objects appear not to exist (e.g. there is no round square, yet we can entertain the thought that the round square is round), but if intentionality is a relation then this cannot be so—for a relation to hold, its *relata* must exist. One must either give up the notion that intentionality involves a relation (for a time, Brentano thought intentionality was quasi-relational—meaning, roughly, that it involved not a relation, but something relation-like), make sense of the oxymoron that is the claim that there are non-existent objects, or give up the notion that intentionality is a relation that necessarily involves intentional objects (though, without giving up the notion that intentionality does involve a special kind of relation). If one opts for either of the two initial options, one endorses *the object theory of intentionality* (Sajama and Kampinen, 1987).

The analytic tradition has, following Frege and Russell, opted to go with the third option: intentionality involves a subject being in some relation to a content, and not an object. Contents can roughly be thought of as descriptions of states-of-affairs. This theory is called *the content theory of intentionality* (Sajama and Kampinen, 1987). We thus start our discussion with the following picture: intentionality is that which warrants content attributions, where contents are construed as accuracy or satisfaction conditions (Seager and Bourget, 2007). This view is quite flexible and is tacitly or explicitly endorsed by multiple authors of interest (in particular, Horgan, 2013 and Horgan and Tienson, 2002). However, we shall see that phenomenal intentionality blurs the neat distinction between object and content theories of intentionality. There are two general questions that arise in connection to content attributions: (i) what contents can properly be attributed and (ii) when, and what things can properly be the subjects of content attributions. A central concept to the latter question was brought to the fore by Searle (1980): the notion of intrinsic, or original, intentionality, which we discuss below (see 'Horgan's Morph Sequence Argument').

II. Towards Phenomenal Intentionality: Phenomenology and Intentionality Together

Phenomenology is a part of experience. The phenomenology of an experience is the way the experience appears to its experiencer. It is the first-person, introspectible aspect of an experience of which its experiencer is aware. An experience is an event with a patient—its experiencer—and always² has a phenomenology. An experiencer's phenomenology is not identical to or completely accounted for by some or all of the experiencer's capacity to differentially respond to stimuli; one can seemingly have a discriminative capacity while apparently lacking corresponding phenomenology—there apparently needs to be something more (Siewert, 2000).

It was common for analytic philosophers to hold that phenomenality and intentionality could be understood independently (Seager and Bourget, 2007; Pitt, 2004; Horgan and Tienson, 2002; Lycan, 2008), a view that Horgan and Tienson (2002) call *separatism*. This trend seems to have been losing ground for the past twenty years. Phenomenal intentionality is a decidedly anti-separatist thesis, but it is not the only one; there are multiple extant alternatives (see Seager and Bourget, 2007).

There are multiple ways to deny *separatism*. One way to do this is to claim that phenomenology is intentional, whether in part or in its entirety. Another approach is to claim that some paradigmatically intentional phenomena are also phenomenal. Of these two alternatives, the first is more popular and arguably less controversial. Phenomenal intentionality itself is a remarkably strong anti-separatist notion. It draws on and enhances both of the claims above. This is why Horgan and Tienson (2002) discuss both “the intentionality of phenomenology” and “the phenomenology of intentionality” prior to their defense of phenomenal intentionality. We follow the same pattern.

II.i Phenomenology Viewed as an Intentional Phenomenon.

Proponents of phenomenal intentionality approach the intentionality of phenomenology from a first-person perspective: they point to a kind of intentionality inherent in our experiences. For instance, Horgan and Tienson (2002) point out that whenever we have a phenomenal

² Or sometimes. It depends on your inclinations. See, for example, Byrne (2004).

experience, it is always *about* something (i.e. objects whose properties we experience as a unity and through time), and that without this aboutness there would not be any phenomenal experience to talk about. Intentionality is therefore inseparable from the experience's phenomenal character. They also point out that the intentionality of phenomenology has often been missed. When philosophers talk about phenomenal experiences it is often in the form of a scenario wherein the object of the experience is ignored in service of highlighting the subjective aspect of the experience in question. As such, it comes as no surprise that the intentionality of phenomenology is missed. However, even if it were the case that we were only talking about the “pure” experience, the intentional content inherent in phenomenology would still be there. As Horgan and Tienson put it:

For any experience involving a specific shade of red, one can abstract away from the total experience and focus on the distinctive what-its-like of that shade of red *per se*—a phenomenal aspect of this total experience that it has in common with innumerable other total experiences that differ in the perceived location of the experienced red or in the shape of the red surface, etc. *But even considered in isolation from any total visual-experiential state, the what-it's-like of experiencing red is already intentional, because it involves red as the intentional object of one's experience.* Again, redness is not experienced as an introspectible property of one's own experiential state, but rather as a property of visually presented objects. (pg. 521; emphasis added).

II.ii Intentionality Viewed as a Phenomenal Phenomenon

Traditionally, thoughts, beliefs, desires, hopes, fears, and the like—the attitudes—are thought not to have an accompanying phenomenology. Phenomenal intentionality is associated with an unorthodox notion concerning this matter—cognitive phenomenology. In a review of the topic, Lycan (2008) traces the thesis back to Goldman (1993), though

Kriegel (2013) claims arguments in its favor go, at least in form, as far back as Chisholm (1957). According to Horgan and Tienson:

Intentional states have a character, and this phenomenal character is precisely the what-it-is-like of experiencing a specific propositional-attitude type vis-a-vis a specific intentional content. Change either the attitude-type (believing, desiring, wondering, hoping, etc.) or the particular intentional content, and the phenomenal character thereby changes too. Eliminate the intentional state, and the phenomenal character is thereby eliminated too. This particular phenomenal character could not be present in experience in the absence of that intentional state itself. (pg. 522)

Having a particular conscious belief about a rose (e.g. that you enjoy looking at it) inherently comes with the phenomenal experience of the particular intentional content. In addition, there is also something that it is like to have a belief about the rose in the first place. David Pitt (2004) defends a similar view:

Each type of conscious thought—each state of consciously thinking that *p*, for all thinkable contents *p*—has a proprietary, distinctive, individuative phenomenology. (pg. 5)

To give us a glimpse at what this phenomenology is, Pitt asks us to consider the following syntactically challenging sentence:

- (1) The boy the man the girl saw chased fled.

Pay attention to what your reading of (1) is like for you, and call it φ . Pitt tells us that (1) means the same thing as:

- (1') The boy, who was chased by the man that the girl saw, fled.

Now, return to (1), and again pay attention to what your reading is like for you, call it ψ . Pitt claims there is a phenomenal difference between φ and ψ , and further that the only difference can be your lack of the proprietary, distinctive and individuative phenomenology that (1) in φ .

According to Pitt, when you have a conscious thought that (1), not only are you related to some content—namely the proposition that (1)—but you also have some unique corresponding cognitive phenomenology—a proprietary, distinctive, individuative phenomenology that (1). Apart from phenomenological demonstrations of the sort above, Pitt offers the following argument:

Normally [...] one is able consciously, introspectively, and non-inferentially (henceforth "Immediately") to do three distinct (but closely related) things: (a) to distinguish one's occurrent conscious thoughts from one's other occurrent conscious mental states; (b) to distinguish one's occurrent conscious thoughts each from the others and (c) to identify each one of one's occurrent conscious thoughts as the thought it is (i.e., as having the *content* it does). But [...], one would not be able to do these things unless each (type of) occurrent conscious thought had a phenomenology that is (1) different from that of any other type of conscious mental state (proprietary), (2) different from that of any other type of conscious thought (distinct), and (3) constitutive of its (representational) content (individuative). (pg. 7; emphasis in original)

By 'representational content,' Pitt is referring to "those properties in virtue of which [a thought] represents (expresses) the proposition it does" (pg. 6). This allows Pitt to remain neutral with regards to any specific content theory of intentionality.

III. Phenomenal Intentionality

In addition to the intentionality of phenomenal experience and the

phenomenality of intentional attitudes and contents, Horgan and Tienson (2002) argue for a kind of intentionality that is entirely phenomenal—Phenomenal Intentionality (henceforth PI). We break PI up into two broad theses: the existence thesis and the primacy thesis. PI has been analyzed into various other theses (see Kriegel, 2013), but we think that, in the interest of generality and parsimony, this is the best way to look at it (Mendelovici & Bourget, 2014, provide a similar analysis). Almost every proponent of PI endorses that it exists (existence thesis) *and* that there is something special about it, something that renders the contemporary paradigm unable to account for PI (primacy thesis).

III.i The Existence Thesis

Horgan and Tienson argue for the existence of PI via their *Phenomenal Duplicates* thought experiment. Consider a being whose phenomenology is identical to yours. Other than its phenomenology, we do not know anything else about this being (its environment, physiology etc.). For all we know this being could be some sort of Cartesian ghost or a brain in a vat. The point is that, by creating this ‘veil of ignorance’, the thought experiment can filter out all of the factors other than the phenomenology (external factors, especially).

Now recall that phenomenal experiences (e.g. color-experiences) have intentional content that is inseparable from their phenomenal character. If this is right, then it follows that your phenomenal duplicate will be presented a world in exactly the same way as it is to you. Now imagine looking at a picture hanging crooked on a wall. Your phenomenal duplicate will have the same experience with the same phenomenal content—the same rich set of phenomenal characters that make up the phenomenal experience of looking at some picture hanging crooked on a wall. Whether or not your phenomenal duplicate’s experience is veridical depends on the environment it is in, for example, whether there actually is a crooked picture, or whether your duplicate is a brain in a vat. However, Horgan and Tienson (2002) contend that it is the phenomenal content which you and your duplicate have which determines the accuracy conditions of the phenomenal experience that you are both having independently of the environment. These accuracy conditions are provided by a special kind of content, namely,

phenomenal intentional content. That is not to say that your experience has no dependency on the environment, but only that the accuracy conditions of your phenomenal intentional content depend only on the contents of your experience. We can vary the environment all we want; so long as the phenomenology remains constant, the phenomenal intentional content will also remain constant.

It is not quite clear whether or not phenomenal content can stay constant while the environment is varied—phenomenal externalists argue that it cannot. As such, one could dismiss PI on the grounds that the thought experiment does not work on principle. However, it is not clear from Horgan and Tienson's (2002) use of the term 'content' that they are using it in the orthodox sense. If anything, the phenomenal duplicates thought experiment points toward an interpretation of *phenomenal intentional content* as being a kind of phenomenally accessible feature. Let us call this interpretation the *phenomenal feature* interpretation.

One might object that the phenomenal feature interpretation is explicitly denied by Horgan and Tienson (2002) when they say that “the sensory-phenomenal experience, by itself, determines conditions of accuracy: i.e., a class of ways the environment must be in order for the experience to be accurate”. However, the main motivation for the phenomenal feature interpretation comes from the quotation in Section 2.1, where Horgan and Tienson (2002) actually seem to endorse it: “the what-it's-like of experiencing red is already intentional, because it involves red as the intentional object of one's experience” (pg. 521). Further, as noted above, several passages in the text suggest that the accuracy conditions determined by phenomenal intentional content are ‘phenomenologically determined’ and emphasize the fact that these conditions are not sensitive to the environment for their individuation. This is strange, as Horgan and Tienson (2002) also do not deny externalism. Instead, they distinguish *narrow* and *wide truth conditions*, a move that leaves us with serious doubts as to whether the two are the same kind of thing.

Now, we have outlined two broad ways of interpreting “phenomenal intentional content”. By the first, phenomenal intentional content is a kind of content in the ordinary sense of the term. By the

second, phenomenal intentional content is a kind of phenomenal property. If we adopt the first interpretation, the proponent of PI has to argue for the narrowness of phenomenal content—something we do not endorse. However, we do not have any immediate grievances with the second interpretation. As such, if the second interpretation is right, then we believe that the existence of PI is, at the very least, plausible.

III.ii The Primacy Thesis

The primacy thesis takes things to another level; not only does PI exist, but traditional theories of intentionality are apparently unable to account for it. It will be useful to characterize which theories of intentionality proponents of PI find inadequate—to characterize just which theories these “traditional” theories are. Unfortunately, there does not seem to be a clear criterion picking out these theories. We can still, however, carve out the space approximately.

Uriah Kriegel (2013) identifies these theories as those which subscribe to what he calls the Naturalist-Externalist Research Program. In this program, intentionality is typically “construed as involving in its core a type of tracking relation, whereby internal states occur sensitively to the presence of specific external conditions” (pg. 1). Causal, covariational, teleosemantic and learning-based theories of intentionality all fall under the purview of this characterization (Horgan and Tienson, 2002).³ Externalist theories of intentionality also seem to be targeted, if only because they tend to be of one of the types above. Some arguments for primacy directly target externalism, as we shall see. In this section, we consider and reject two main arguments for the primacy thesis, the argument from the narrowness of phenomenology and Horgan's (2013) Morph Sequence Argument.

III.iii The Argument from Narrowness.

One way to show that PI cannot be accounted for by other theories of intentionality is to show that phenomenology is narrow. For the purposes of this discussion, a property is narrow if and only if it supervenes only on the local properties of the thing that instantiates it (Tye, 2007)—micro-physical duplicates share their narrow physical properties and

³ Such theories would be the topic of Stitch and Warfield (1994).

differ in their wide properties, or so the story goes. In other words, phenomenology is narrow if and only if internal properties (of the brain in our case) wholly constitute it. If phenomenology is narrow, then it follows that PI is also narrow. After all, if there is a kind of intentional content that is completely determined by phenomenology, and phenomenology is narrow, then it follows that there is a kind of intentional content that is completely narrow.

To motivate their claim to the narrowness of phenomenology, Horgan and Tienson (2002) make a distinction between *causal* and *constitutive* roles in regards to phenomenology. They do not deny that the environment *causes* one to have certain phenomenological experiences; however, that does not thereby mean that the environment partly *constitutes* one's phenomenology. Furthermore, regardless of the fact that environmental stimuli cause phenomenal experiences, the connection is not direct. The stimuli must first be picked up by peripheral sensory processes (eyes, skin, etc.), then transferred to the brain via neural pathways, then processed in the appropriate way by the brain, and then, finally, the phenomenology emerges.

Horgan and Tienson's (2002) claim can be viewed as placing a limit on the physical states upon which phenomenal states can supervene. They say that so long as one's sensory receptors exhibit certain patterns of activation, we can, given some other knowledge about one's brain, tell which phenomenal states one will be in regardless of the factors causing the patterns of activation. In that sense, the environment can be said to be causing one's phenomenology, though it plays no constitutive role in it. Intuitive as this picture of phenomenology may be, it is a controversial one. Phenomenal externalists deny this picture with a variety of thought experiments that exhibit brain structures that are radically decoupled from the environment (Tye, 2007).

However, Horgan and Tienson (2002) do not so much establish the narrowness of phenomenology as they advertise its consequences for PI. Thus, if one is a non-believer about the narrowness of phenomenology, one will also remain unconvinced by the argument from narrowness for the primacy of PI. We are not convinced that phenomenology is narrow. As such, as long Horgan and Tienson do not

meet externalist objections to the narrowness of phenomenology, there will be no way forward for this line of argument.

III.iv Horgan's Morph Sequence Argument

In contrast to the argument from narrowness, Horgan (2013) provides a direct argument for the primacy of PI. He calls it the Morph Sequence Argument (henceforth MSA). The MSA is based on Searle's (1980) Chinese Room thought experiment. Briefly, imagine a guy who is a non-Chinese speaker and is tasked with responding to queries, all in writing, in an isolated room. The queries and responses are sequences of Chinese characters. The guy is never given any information about the meanings of the queries or responses, he is just given very detailed procedural instructions for matching queries to responses in another language he understands. However, it turns out that the queries are actually well-formed Chinese utterances, and the responses are well-formed felicitous Chinese language responses to those utterances. Searle's contention is that, despite the responses' being intelligible to those outside the room, the guy "understands no Chinese" and therefore has no Chinese language intentionality.

Searle's point is that functionalist accounts of intentionality are insufficient in explaining original intentionality (henceforth OI); that is to say, the intentionality that some mental states, such as beliefs, desires and so on, have intrinsically.⁴ Horgan tries to harness the intuitions mobilized by the Chinese Room in the MSA to drive home his point—the primacy of PI. The MSA, as the name suggests, follows a series of scenarios which teases our intuitions, little by little, to lead to the conclusion that what is missing in the Chinese Room is PI; or in other

⁴ Some things, such as this text, are intentional—they afford warranted content attributions—but only, the claim goes, insofar as they are artifacts which warrant interpretations. But an interpretation is an intentional state: there must be an interpreter to whom the interpretation can be attributed. Thus, such artifacts have intentionality in virtue of more basic sorts of intentionality. Searle, and many others, believe that this chain of interdependence bottoms out in certain distinctive states which are intentional in and of themselves. Such states are intrinsically, or originally, intentional and creatures that have them have intrinsic, or original intentionality.

words, PI is OI.

The first scenario is the Chinese Room in its original form. Horgan argues that we all agree that the guy understands no Chinese.

The second scenario provides the guy with a compact symbol manipulator appended to his brain; Horgan calls this the “monitoring/processing/stimulation” (MPS) device. The MPS device does exactly what the guy did in the room. The difference is that it is faster and automatic. Essentially, the MPS device monitors the guy's visual input as he reads the symbols. It then manipulates the symbols based on the same procedural instructions as in the first scenario, after which it sends various neural signals to the guy's brain such that he spontaneously writes down certain symbols on a piece of paper. Horgan argues that we all agree that the guy still understands no Chinese.

The third scenario generalizes the MPS's function to a new modality: speech. As with the guy's visual input, the MPS device now tracks the guy's auditory input. The MPS device has the ability to recognize Chinese speech sounds and map them onto the various symbols that characterize them in written language. As such, the MPS device can manipulate auditory input in the same fashion as visual input. It also sends various other neural signals to the guy's brain which cause him to spontaneously make meaningless-to-him vocal noises. Unbeknownst to him, those vocal noises are actually spoken Chinese. Horgan argues that we all still agree that the guy understands no Chinese.

The fourth scenario gives the guy a telescope with which to see the people outside the room. He can see the people talking to him, talking to each other, and writing down symbols on pieces of paper. Unfortunately, he has a strange case of anterograde amnesia which prevents him from remembering anything after thirty seconds. This prevents him from learning Chinese via behaviors, lip movements, etc. that he is now observing. The MPS is still attached and keeps its updated function from the third scenario. Horgan argues that we all still agree that the guy understands no Chinese.

The fifth scenario gives the MPS a few upgrades. First, the MPS's monitoring and actuating abilities are generalized to include all of the guy's senses, to reflect his mental states and to produce correct body

language. The second upgrade gives the MPS the power to generate in the guy (non-cognitive) sensory images, emotions, and any other non-cognitive phenomenology that normally arises in a normal Chinese speaker. The third upgrade makes it such that the MPS device prevents the guy from having the sort of mental state one would normally have, given the situation he is in (e.g. confusion, fear, etc.). Additionally, the guy now interacts with the Chinese population outside the room. He converses, laughs, shakes hands etc. with them in such a way that an ordinary Chinese speaking person would not be able to tell him apart from any other Chinese speaking person. Horgan argues that we all still agree that the guy understands no Chinese.

The thrust of the argument is that, without the right kind of phenomenology, one cannot have the intentional content associated with that kind of phenomenology. In the guy's case, he does not have the intentional content associated with the Chinese language utterances he emits because he has no Chinese language understanding phenomenology. A puzzling question is what exactly this Chinese language phenomenology is supposed to be like. We think that this phenomenology should allow one to immediately (in Pitt's sense) (i) distinguish any utterance of Chinese from any other utterance of Chinese, (ii) distinguish any utterance of Chinese from any other utterance (in any other language) that has the same sense and (iii) to identify any utterance as having the sense that it does. We will not discuss whether or not such phenomenology exists at all; however, without a conception of Chinese language understanding phenomenology similar to the one above, the argument will not get off the ground. In other words, the MSA shows, in the best case, that PI is necessary for intentionality. If PI were OI, it would be sufficient for intentionality; but this, the MSA does not show.

One can reply that the MSA coupled with the argument from narrowness is sufficient to show that PI is OI, but this still does not work. To see this, one can look at the MSA as the second half of a double dissociation experiment between PI and intentionality as construed traditionally. A double dissociation is a common experimental design used to demonstrate that two or more factors are independent of one another. Consider Broca's and Wernicke's areas in the brain. Roughly,

Broca's area functions as the language comprehension area, while Wernicke's area functions as the language production area.⁵ We know this because people with lesions in their Broca's area lack the ability to produce fluent speech while retaining the ability to understand speech, while people with lesions in their Wernicke's area have been shown to produce fluent but nonsensical speech (Van Orden, Pennington & Stone, 2001).

The first half of the experiment is just the narrowness of PI: one can have phenomenal intentional content without in any way being in the right kind of relationship to the world, thus not having the associated content (in the traditional sense). The MSA shows that one can be in the right kinds of external relationship to the world to count as having some content (in the traditional sense) without having the associated phenomenal intentional content. Thus, the two are independent. But, there is still a missing premise, namely, the premise that is supposed to show how PI is more basic than intentionality as construed traditionally. Horgan seems to believe that, because PI is what is needed for language understanding to occur, and that OI is what was missing in Searle's original argument, it follows that PI is OI. However, this is only one possibility. If the scenario was inverted, and PI was there but intentionality as construed traditionally was always missing, would it follow that the guy understands Chinese? We think not. It could easily be the case that both PI and intentionality as construed traditionally are required for OI—that each is necessary but neither is sufficient on their own.

IV. Conclusion

It is not clear whether the primacy of PI is tenable, as it depends on one's commitments with regards to internalism. Furthermore, the idea that PI should be OI seems to be intuitive, but also not very clearly supported. To show that something is OI, one must show that having it is sufficient for having intentionality, but it is not clear how anyone would go about

⁵ In case it was not clear, this is a *very* simplistic characterization of Broca's and Wernicke's area.

doing this. In any case, Horgan and Tienson (2002) as well as Horgan (2013) do not do this. While we focused on the works of Horgan and Tienson, any defender of PI faces the same challenges in so far as they subscribe to both the existence and primacy theses.

However, we should note that what we find genuinely novel about PI is the phenomenal feature interpretation, for it seems to define an approach to intentionality that is genuinely different from the traditional paradigm. But, we should also add the caveat that it does not necessarily contradict the traditional paradigm. The phenomenal feature interpretation seems able to capture the spirit of the object theory without the drawbacks.

Works Cited

Byrne, A. (2004). What phenomenal consciousness is like. *Advances in Consciousness Research*, 56, 203-226.

Chisholm, R. (1957). *Perceiving: A Philosophical Study*. Ithaca : Cornell University Press.

Goldman, A. (1993). Psychology of Folk Psychology. *Behavioral and Brain Sciences*, 16, 15–28.

Fodor, J. (1994). Fodor's Guide to Mental Representation: The Intelligent Auntie's Vade-Mecum. In Stich, S. and Warfield, T., eds. *Mental Representations: A Reader*, 9–33 Blackwell.

Horgan, T. (2013). Original Intentionality is Phenomenal Intentionality. *The Monist*, 96(2), 232–251.

Horgan, T., and Tienson, J. (2002). The Intentionality of Phenomenology and the Phenomenology of Intentionality. In Chalmers, D. (ed.), *Philosophy of Mind: Contemporary and Classical Readings*, 520 – 532. See also:
<http://www.u.arizona.edu/~thorgan/papers/mind/IPandPI.htm>

Kriegel, U. (2013). The phenomenal intentionality research program. In Uriah Kriegel (ed.) *Phenomenal Intentionality*. Oxford University Press.

Lycan, W. G. (2008). *Phenomenal intentionalities*. *American Philosophical Quarterly*, 233-252.

Mendelovici, A., Bourget, D. (2014). Naturalizing Intentionality: Tracking Theories vs. Phenomenal Intentionality Theories. *Philosophy Campus*, 9(5), 325 - 337.

Pitt, D. (2004). The Phenomenology of Cognition Or What Is It Like to Think That P?. *Philosophy and Phenomenological Research*, 69(1), 1 – 36.

Sajama and Kampinen (1987). *A Historical Introduction to Phenomenology*. New York: Croom Helm.

Searle, J. (1980). Minds, brains, and programs. *The Behavioral and Brain Sciences*, 3, 417 – 457.

Seager, W. and Bourget, D. (2007). Representationalism about consciousness. In Velmans and Schneider (ed.), *The Blackwell*

Companion to Consciousness. Blackwell. Retrieved from
<http://www.dbourget.com/RAPC.pdf>

Siewert, C. (2000). Precis of The Significance of Consciousness.
PSYCHE, 6, 12.

Stich, S., and Warfield, T., eds. (1994). *Mental Representations: A Reader*. Blackwell.

Tye, M. (2007). New troubles for the qualia freak. In McLaughlin & Cohen (eds.), *Contemporary Debates in Philosophy of Mind*. Blackwell.

Van Orden, G. C., Pennington, B. F., and Stone, G. O. (2001). What do double dissociations prove?. *Cognitive Science*, 25(1), 111-172.

The Telltale Footprint: Doctored Evidence and Epistemic Value

Ryan Martins

In his paper “Getting Told and Being Believed,” Richard Moran argues for what he calls the Assurance View of Testimony. In order to do so, he first attempts to clear the ground from a prominent family of views which take testimony as evidence. The main thrust of his critique is that these evidential views of testimony seem to be analogous to cases of doctored forensic evidence. He challenges the evidentialist to show how testimony’s deliberate presentation could in any way add, rather than detract from, the testimony’s evidential worth, as we take it to do with other forms of evidence.

In order to put Moran’s objection in context, it is important to note a few relatively major considerations that are affected by the success or failure of his objection. Remember that Moran ultimately seeks to put forward his own view, the Assurance View, which moves away from evidence, and instead focuses on the exchange between a speaker and a hearer, and the assurance that one receives from the other. In order to do this, he first purports to expose a flaw in the competing evidentialist picture by way of his unfavorable analogy. The first thing to notice becomes evident once one considers the role testimony plays in knowledge. Testimony, along with sense perception, is one of the most pervasive sources of our knowledge. If Moran is correct in his assessment of the evidential view, then it seems like testimony functions as an epistemically irresponsible source of knowledge – a clearly undesirable result. Second, if Moran is right and his analogy holds, then the evidentialist’s picture is critically flawed – it is committed to conceiving of all testimonial evidence as doctored. Since his own view forgoes the evidential model in favour of one that highlights the relationship between the speaker and hearer, Moran has a strong reason to accept the Assurance View over existing evidential accounts. If,

however, the evidentialist can defend herself from his objection, then Moran's view loses this advantage, and he must find some other reason for us to prefer his Assurance View over the evidentialist accounts.

In this paper I seek to address Moran's concerns on behalf of the evidentialist by outlining two distinct, but compatible, responses to Moran's objection. The first involves accepting Moran's analogy, but distinguishing between general intention and the specific motivation to deceive or mislead another. The second claims that there is a key disanalogy in our default assumptions about the intentionality of testimony and forensic evidence, and that this disanalogy is a sufficient basis for treating them differently. With these two responses in hand, I conclude that the evidentialist is well prepared to rise to Moran's challenge and defend the notion of testimony as essentially evidential.

This paper will be divided into the following sections: first, I provide an introduction to some relevant concepts. In the second, I discuss Moran's analogy, and evaluate how well the concept of forensic evidence captures the evidentialist's view of testimony. In the third, I outline what it is about this analogy that creates a problem for evidential views of testimony, commenting in particular about the concepts of dependence and intentionality. In the fourth section, I discuss two possible responses available to the evidentialist, and comment on their effectiveness as means of combating the problems Moran's analogy has raised. And finally, in the fifth, I offer concluding remarks.

I. Setting Up

Most work in the epistemology of testimony has centered on what may be called evidential accounts of testimony. Testimony, according to these views, provides those who hear it with evidence for a particular belief.¹ It is exactly these accounts—those that rely on testimony as being evidential—that Moran seeks to counter with his own Assurance View. One of the methods he uses to do so is an argument from analogy. He

¹ I sometimes speak of “testimonial evidence” for ease of exposition. It is to be noted that I am referring to testimony reconstructed on an evidential account, and not necessarily testimonial evidence of the type you would find in a courtroom or similar setting.

seeks to show that the evidence provided by testimony is similar to doctored forensic evidence—physical pieces of evidence of the kind we would find at the scene of a crime—and that certain characteristics of this evidence—namely, its reliance on deliberateness and intentionality—seem to pose a problem for adherents of the evidential view.

Evidential views maintain that testimony is a type of evidence for the considerations at hand. An exchange between a speaker and hearer, then, is one in which the speech act of the speaker may be taken as evidence for some proposition. When I tell you “it is snowing outside”, you may take that assertion (at least) as evidence of the proposition, “I believe it is snowing outside.” It is this relationship with evidence that Moran thinks opens up these accounts to problems like the one his analogy raises. By treating testimony as evidence, evidential views of testimony must address a critical issue, which can be illustrated by an analogy with forensic evidence.

Before I get to the analogy, it should be noted that the concept of forensic evidence in play is one in which it is completely divorced from testimony. One should think of forensic evidence as being found ‘at the scene of the crime’; you are actually *seeing* the footprint on the ground, actually *smelling* the smoke in the air. Forensic evidence in this sense is not reliant on testimony, but instead on direct perception. You are discovering this evidence yourself; it is not being presented to you through some intermediary agent, like a lawyer presenting it to a jury at a trial.²

II. The Analogy

Moran maintains that if we treat testimony as evidence, it must work in much the same way as forensic evidence, but with one key disadvantage. It must not be treated as just any forensic evidence, but *doctored* forensic evidence, and this, Moran thinks, poses a serious problem for evidential

² This is not, of course, to suggest that forensic evidence *cannot* be presented by a speaker; the example I give of a lawyer in a trial is one of many such instances where this does happen. However, since we are comparing forensic evidence to testimony, it would complicate things unnecessarily to have this forensic evidence filtered through testimony.

views of testimony. Doctored evidence is to be taken as the kind of forensic evidence that has been either altered in some way from its original form, or fabricated entirely. There are two main components which are found in the case of doctored forensic evidence that need to be present in order for the analogy to hold in the case of testimonial exchange. The first is a deliberative component, which concerns the act by which the evidence is presented. The second is an intentional component, which concerns the motive of that action. The deliberative component requires that the act by which the evidence is presented to the other party be a deliberate one. The evidence must, in this sense, be presented to the other party on purpose. In addition to being done deliberately, the evidence must be also presented with a certain motive. The intentional component requires that the one presenting the evidence must have the intention of leading the other party to a certain belief. The content of this belief matters not, as will be discussed in section IV. It must simply be the case that the evidence is being presented in order to lead the other party to *some* belief. For the analogy between doctored forensic evidence and testimony to hold, then, it must be the case, that in a case of testimonial exchange, we find that the speaker deliberately is presenting the hearer with evidence with the intention to lead them to a certain belief.

If we take testimony to be evidential, then it is clear that these same components can be found in testimonial exchange. Let us use the example of a testimonial exchange mentioned above in which I told you that it was snowing outside. In this case, one in which a speaker is telling a hearer something, the assertion is indeed made deliberately—the words did not slip accidentally from the speaker's lips. It is difficult to see how an assertion can be anything but deliberate, except in exceptional cases like talking in your sleep. Such an assertion also manages to satisfy the requirements of the intentional component—it is done with the intention of leading the hearer to a particular belief. When I tell you it is snowing, I am informing you (at the very least) that I believe that it is snowing outside. I do this with the intention of informing you of my beliefs; I am guiding you to discover what I believe. Given these facts, it is clear that testimony can satisfy the requirements of both the deliberative and

intentional components that are found in doctored forensic evidence, and so Moran's analogy holds. In both cases of doctored forensic evidence and testimonial exchanges one person is deliberately presenting evidence with the intention of having his or her audience form a certain belief.

III. The Problem

It has been established that testimony, on the evidential view, and what we have been calling doctored forensic evidence, are analogous, at least with respect to the two crucial components mentioned above. Now, we need to inquire into why this analogy may pose a problem for evidential accounts of testimony.

The problem, as Moran sees it, stems from how we tend to treat forensic evidence after discovering it has been placed or presented intentionally. Ordinarily, when one comes across forensic evidence, like a footprint in the snow or a discarded cigarette, there is little issue with accepting it as such. However, upon discovering that the evidence was left there intentionally, we immediately discredit its evidential worth. It is not, as Moran puts it, "*better* evidence, or even just as good, but instead like something fraudulent, or tainted evidence."³ If this is the case, and the analogy holds, it seems we must consider acts of testimony as a kind of *tainted* evidence for the speaker's beliefs.

It is recognition of the deliberate and intentional character of forensic evidence that discredits it in our eyes. Upon recognition of this intent, the evidence ceases to be a clue helping us to discover something, and instead strikes us more as an attempt to mislead. If, indeed, the problem does stem from recognition of intentionality, this is a problem that will especially affect testimony, as there is no way to divorce intentionality and testimony. Evidence of the forensic kind can be left behind unintentionally, like the carelessly discarded cigarette butt, and it is actually this kind of unintentionally left behind evidence that we prefer in forensic cases. There is no necessary connection between intentionality in this sense and forensic evidence. Testimonial evidence, however, is quite the opposite. One cannot separate testimony from its presentation. As Ross notes, a photograph has evidential value whether

³ Richard Moran, "Getting Told and Being Believed," *Philosophers' Imprint*, Vol. 5, No. 5 (Aug., 2005), 6. Emphasis in the original.

or not someone presents it to you as such; the same, however, is not true of testimony.⁴ There is no testimony without presentation, no way to have someone stumble upon some piece of testimonial evidence like they would a footprint. Testimony, then, does not appear to have evidential value independent of its presentation, unlike forensic evidence, which maintains its evidential value regardless of its being presented.

Given that testimony cannot be divorced from intention, it appears that testimony is particularly vulnerable to any criticism that targets its intentionality. As Moran puts it, “if we are considering speech as evidence, we will have to eventually face the question of how recognition of its intentional character could *enhance* rather than detract from its epistemic value for an audience.”⁵ If intentionality detracts from the value of evidence, and testimony cannot be divorced from intentionality, it seems then, that testimony is a rather poor form of evidence. It is this issue that presents a serious challenge to the force of evidential theories of testimony. As such a pervasive source of knowledge—one that we rely on almost more than any other—we would not want a theory of testimony that concludes that testimony is a poor and unnecessarily risky form of learning a speaker’s beliefs. As Moran notes, if we conceive of evidential testimony in this way, then there is no reason to privilege testimony as a source of knowledge. In fact, we should do the opposite. Any other method of learning the speaker’s beliefs would be preferable, especially those that do not depend on the intentionality of the speaker, like hypnosis or spying. If what he says is true, Moran concludes, testimony is just a “very possibly misleading” way of learning a speaker’s beliefs.⁶

Not only should testimonial evidence be regarded as a rather poor kind of evidence due to its inherent connection with intentionality—which seemingly taints such evidence—but also because we are subject to the will of the speaker in using testimony as evidence (that is, we are dependent on the speaker). By using testimony to gain access to the

⁴ Angus Ross, “Why Do We Believe What We Are Told?,” *Ratio*, Vol. 28, No. 1 (1986), 72.

⁵ Moran, “Getting Told and Being Believed,” 6.

⁶ *Ibid.*

speaker's beliefs, we are exposing ourselves to unnecessary epistemic risk. We are at the mercy of the speaker to truthfully inform us of their belief, and it is not at all out of the ordinary to find out that they have intentionally misled us. The same is not true of forensic evidence. When we see a photograph, we are not dependent on it to reveal the truth to us in the same way. We might fail to infer the correct conclusion from the photograph, but that is in no way the fault of the photograph itself. If testimony is to be taken as a kind of evidence, it is a kind where we are taking on additional epistemic risk due to our need to rely on the intentions of the speaker.

Moran structures his analogy between testimonial and doctored forensic evidence in order to point out a problem that arises for the evidentialist. Testimony is unavoidably connected with intentionality, yet intentionality *detracts* from the evidential value of a piece of evidence. By taking testimony merely as evidence, we are exposing ourselves to great epistemic risk, as we are wholly dependent on the speaker to provide us with proper evidence. Cases where they fail to do so, such as lying, are not at all out of the ordinary. If this is the case, then testimonial evidence functions as an extremely poor way of learning a speaker's beliefs, an undesirable conclusion for any serious account of testimony.

IV. Two Responses

Confronted with these problems, there are several responses available to the evidentialist, two of which I will discuss here. First, however, it is important to note that the evidentialist *cannot* avoid the issues raised by Moran by suggesting that we move from thinking of the testimony as evidence of the speaker's belief, and instead consider it to be evidence of the truth. The problems with the evidential picture, pointed out by Moran's analogy, arise from the notions of intentionality, dependence, and deliberateness. These elements are all still present in the case where the speakers are using their testimony as evidence of the truth, rather than as evidence of their beliefs. The problem has to do with the intentionality of the presentation of the evidence, and the hearer's dependence on the speaker, both of which are not at all affected by the fact that the speaker is pointing you to the truth rather than to their own beliefs. Evidence is evidence; what it is evidence *of* plays no role as to its

status as evidence. With this addressed, I wish to evaluate what I see as some of the stronger responses available to the evidentialist in order to respond to the problems raised by Moran's analogy.

It seems plausible for the evidentialist to agree with Moran's analysis, but to question the force of his broader argument that recognition of intent detracts from the testimony's evidential value. They may accept that testimonial evidence is, at least with respect to the two components of intentionality and deliberateness, akin to doctored forensic evidence. The evidentialist might also conclude that we are exposing ourselves to epistemic risk when relying on testimonial evidence. They may even agree that because of its necessary connection with intentionality, it is difficult to see how we can view testimony as anything but doctored evidence. The evidentialist however, can then point out that Moran never distinguishes between mere general intentionality, and the more specific intent to deceive or mislead.

Moran is correct in thinking that when we come across evidence, and later find out it has been placed there deliberately, we discredit its evidential worth. Moran seems to think it is our recognition of the intentional character of the evidence that causes us to discredit it as good evidence, and so, we are forced to conclude the same when recognizing this intentionality in testimony. But why think this is the case? There are any number of everyday cases in which someone is deliberately presented with testimonial evidence in order to lead them to a certain belief, but we do not somehow conclude that their testimony functions as poor evidence. Upon arriving in a new city, you ask someone the directions to your hotel. They respond by saying "It is two blocks East". They have presented you with this assertion deliberately, as they were responding to your question, and they have done so with the intent of having you form a certain belief, namely, that the hotel is two blocks east (or at least the belief that they believe this is the case). Should we be suspect of this assertion merely because it has been done deliberately? Does her assertion, taken as evidence, strike you as tainted, or fraudulent merely because it was done so with the intention of having us form a certain belief?

Intuitively, the answer to these questions is no. This is because it

is not the recognition of mere general intentionality that causes us to discredit or question the evidence, but the recognition of certain specific intentions, like the intent to mislead or deceive us. The intent to mislead is one of an infinite number of possible intentions available to an agent. Why should we not consider specific intentions when determining if intention detracts from epistemic value instead of lumping all intentions together? In the earlier case where someone is asking for directions, the speaker appears to be acting from the intent to inform. This is why the speaker's assertion (that the hotel is two blocks east) does not strike us as tainted or fraudulent. The problem for the evidentialist is that testimony and intentionality are inseparable, and it is recognition of intentionality that is responsible for discrediting the testimony's evidential value. However, not all intention is *intention to mislead*. In light of all the intentions that agents are capable of having, there is no reason for us to default to thinking that others have unscrupulous ones. It is untenable to think that because agents are capable of misleading us, we should always act as if they are – they are equally capable, and generally more inclined, not to mislead us. If it is only recognition of certain specific intentions like the intent to mislead that is responsible for discrediting testimonial evidence, and not the recognition of the presence of mere general intention, then the fact that testimony is inseparable from intention no longer poses an issue for the evidentialist, as not all testimony is intended to mislead.

So, while we recognize that by offering testimonial evidence, an agent deliberately leads us to form a certain belief, this does not mean that this belief is untrue, or that she is trying to mislead us. One can see why, in the case of forensic evidence, one might think this way: the criminal doctors the evidence in order to get us to believe something false, so that she may avoid detection. She has a large incentive to try and mislead us. But the motives of a criminal, and the motives of most people, in most cases of everyday testimonial exchange, do not align. While it is true that people do lie, deceive, and mislead one another, in most cases they do not have incentive to do so. In fact, under normal conditions, people often have large incentive to not intentionally deceive those around them. The mere fact that the possibility of deceit exists is no reason to be suspect of testimonial evidence as a whole. It is not when

we recognize general intent that we should regard the testimony as tainted or fraudulent, but only when we detect certain types of intentions, like the intent to mislead.

A second strategy available to the evidentialist is to point out that there is one very important disanalogy between testimony considered as evidence, and doctored forensic evidence. Normally, we approach testimonial evidence and forensic evidence with different default assumptions about whether or not they are intentional. When we come across a piece of forensic evidence, like the footprint in the snow, our default assumption is that it was not left there with some specific intent: it is just the result of someone walking past. Upon recognizing that it was left there intentionally, however, we shift away from this default assumption, and to a new position in which we know the evidence was left with specific intent. It is this shift away from our default assumption that leads us to be suspicious of some piece of evidence. The suspicion is caused when we consider the fact that the criminal sought to take advantage of this default assumption in order to pass something off as non-intentional when in fact she had placed it there to deliberately mislead. With forensic evidence, we are vulnerable to the possibility that someone may exploit this default assumption in order to gain some sort of advantage.

Testimony however, has an entirely different default state. Our default position with regards to testimonial evidence is that the testimony is, in fact, intentional. There is no point at which we shift away from this default. We do not, as is the case with forensic evidence, shift from thinking the evidence is not intentional, to recognizing that it is. In a case of testimonial evidence, the fact that the evidence is presented deliberately, and is done so with the intention of leading the hearer to a certain belief, is *out in the open*; both the speaker and hearer are aware of it from the very moment the testimony is presented. We know they are speaking to us intentionally. We know they are presenting evidence deliberately to reveal their beliefs. Since our default position with regards to testimonial evidence is that it is intentional, there is no reason for suspicion upon discovering its intentionality—we knew it was there from the start.

The difference in our default attitudes with regards to intentionality in these two cases is important because it provides us with insight into a key disanalogy between testimony considered as evidence, and what we have been calling doctored forensic evidence. What makes the concept of *doctored* evidence so unappealing is not merely the fact that it functions as poor evidence: evidence may be doctored in such a way as to lead one to a correct conclusion. What strikes us as particularly disagreeable is that doctoring is the result of another person's deliberate interference with the evidence. Not only did they tamper with the evidence, but they attempted to pass it off as genuine, unaltered evidence. The default assumption we make when first coming across forensic evidence leaves the possibility of doctoring evidence wide open. Because we assume that forensic evidence has been left behind unintentionally, we leave room for someone to exploit this assumption to pass off doctored evidence as unaltered. But this default is one that only holds for forensic evidence, not testimony. We approach testimonial evidence knowing fully well that it has been presented deliberately, that it has been filtered through an agent before arriving at us. There is no room for someone to exploit this default in order to pass some piece of testimonial evidence off as non-intentional. We have very different default assumption when we approach forensic evidence than when we approach testimonial evidence. While forensic evidence leaves the door wide open for the exploitation of its default assumption, testimonial evidence prevents the exploitation by assuming the very thing that one who would doctor evidence would seek to hide. This disanalogous feature, combined with the distinction between general and specific intent, provides the evidentialists with cogent responses to defend themselves against Moran's charge.

V. Conclusion

In order to put forward his own reconstruction of testimony Richard Moran must first clear the ground by confronting the dominant family of testimonial views, those known as evidential accounts of testimony. One method he uses to do this is to posit an analogy between testimony, considered as evidence, and doctored forensic evidence, the kind of evidence one might find at the scene of a crime. In the second section of this paper, we demonstrated that this analogy holds, particularly in

regards to the two crucial elements of deliberate presentation and the intent to lead the hearer to a certain belief. In the third part of the paper we explained why Moran's analogy raises a problem for the evidential accounts, commenting on the inseparability of testimony and intentionality, and the idea of doctored or tainted evidence. Finally, in the fourth section of the paper we posited two possible responses available to the evidentialist. The first deals with the idea that recognition of mere general intentionality—which is part of every testimonial exchange—is no reason to discredit testimony and is only recognition of certain specific intentions like the intent to mislead. The second response brings to light an important disanalogy between forensic and testimonial evidence. In forensic evidence, we assume no intentionality, and only later discredit the evidence when we recognize that there was, in fact, some aspect of intentionality. However, we do not start from this assumption with testimonial evidence. Instead, the intentionality is placed out in the open, and both parties are aware of it from the beginning of the testimonial exchange. These considerations show that while Moran's analogy may hold, the problems that he thinks it gives rise to, are problems to which the evidentialist can competently respond..

Works Cited

Moran, Richard. "Getting Told and Being Believed." *Philosophers' Imprint*, Vol. 5, No. 5 (Aug., 2005): 1-29.

Ross, Angus. "Why Do We Believe What We Are Told?" *Ratio*, Vol. 28, No. 1 (1986): 69-88.

Hinging Truth: An Examination of the Truth and Falsity of Wittgensteinian ‘Hinge-Propositions’

Christopher Sullivan

“*Im Anfang war die Tat*”
“In the beginning was the deed”
—Goethe, *Faust I*

When a proposition is asserted, it is asserted within a certain context which determines a set of background assumptions and which makes the asserted proposition meaningful. Before we can understand the assertion, before we can question or doubt it, we must presuppose these background assumptions; they act, following Wittgenstein, as ‘hinges’ on which the meaning of the assertion—and our ability to question and doubt it—turn. To what extent are these hinges themselves susceptible to doubt?

G.E. Moore’s infamously attempted to refute philosophical scepticism by holding up his hand and claiming he “knew” it was there. Although Wittgenstein sided with Moore against scepticism and found traction in Moore’s approach to refuting it, he took issue with Moore’s claim that he “knows” he has hand. Claims such as “here is a hand” (or the more general claim that an external world exists) cannot be known or doubted, Wittgenstein argued: we must take them for granted, and treat them as “hinges” on which our empirical knowledge turns.

The topic of this paper is a couple of crucial problems that arise with such propositions. First, it seems deeply troublesome to say that all ‘hinge-propositions’ cannot be either “true” or “false” and to hold that we cannot say, for example, that it is known (and true) that the Earth has existed for a long time, seems absurd. Second, there is the question of why these ‘hinges’ are referred to as ‘propositions’ at all by Wittgenstein scholars, such as Howard Mounce and Annalisa Coliva, if they indeed

lack the qualities of truth and falsity.¹ If we are to grant them any type of ‘propositional’ status, we risk positing a Wittgenstein that is inconsistent with the *Philosophical Investigations*.

This paper will proceed by briefly laying out Moore’s relevant thoughts, so as to illustrate the emergence of Wittgenstein’s ‘hinges’. Then, I will go over the different types of ‘know’ and ‘knowing’ that Wittgenstein outlines in *On Certainty* that will be crucial to a clear understanding of the issue at hand. Finally, I will suggest that, upon closer examination, Wittgenstein can be interpreted as holding that these hinge-propositions can be separated into two groups: those which can be subjected to the light of *evidentiary* examination and those that remain unaffected by such an exercise, which we *take for granted* in an everyday sense. The former type of hinges are therefore not hinges at all, while the latter types of hinges, those which actually deserve that title, are incompatible with the vocabulary of truth and falsity, in the same way Wittgenstein suggests they are incompatible with that of “knowing”. This means we cannot assess or state them to be true. However, in a move I see to be classic Wittgenstein, this is *not* to say that they are *not* true (or false) but merely that they cannot be said meaningfully at all. Therefore, we must treat these hinges as normative in nature for Wittgenstein, and cease terming them ‘propositions’ because they are not really expressing any proposition at all. In doing the above, I hope to forward a reading of *On Certainty* that presents a Wittgenstein very much in line with his *Philosophical Investigations*.

G.E. Moore’s philosophy can be best classified as commonsensical, as opposed to the idealism of Berkeley, or to the scepticism Descartes uses methodologically in his First Meditation. Simply put, Moore believes that scepticism and idealism are false on grounds of their generating a paradox; namely, that if they were accurate, then it would be impossible to generate essential propositions, including those of the sceptics and idealists themselves. In kind, this is not far removed from an initial reading of Wittgenstein; indeed, Wittgenstein was on side with Moore against scepticism. It was Moore’s earlier paper, “A Defence of Common Sense”, which Wittgenstein found particularly valuable, and it can be seen as generating traction for the reflections which eventually became *On Certainty*. In this paper, Moore argues for a

¹ This bipolarity of truth is the condition of being either true or false, and is constitutive of a proposition. See P.M.S. Hacker, *Wittgenstein’s Place in Twentieth Century Analytic Philosophy* (Cambridge: Blackwell, 1996), 35.

number of truisms, which are, in his consideration, *empirical* propositions—which are known and absolutely true.² Crucial to Moore's arguments are the inconsistencies that arise when the idealist denies the reality of material objects,³ since Moore sees the reason that one can have such thoughts is exactly because one is human, with a body.⁴ Furthermore, in the case of the sceptic, if he denies that there are trees, for example, he is left to explain what exactly he is then speaking of.⁵

It is important to take note of Moore's ideas here because they are similar to Wittgenstein's in their focus on something foundational, and particularly nuanced, and too often discounted. Moore's genius is in recognizing the connection of *inference* of one proposition from another, essentially uncovering the architecture on which our everyday propositions are built.⁶ It is in "Proof of an External World" where Moore stumbles in regard to the vocabulary of "knowing", and where Wittgenstein takes issue.⁷

To understand Wittgenstein's criticism of Moore, a going over of the different senses of "I know" is essential. The early paragraphs of *On Certainty* can be seen as a delineation of two concepts of knowing. It is necessary to highlight the differences of these two concepts of 'know' so as to make the semantic and philosophical distinction between them clearer. Simply, the first sense of 'know' is one that is open to doubt and being mistaken; that is, it is open to objective test. A qualified version of the first sense of knowing, as justified true belief, arises from a careful reading, and this will be addressed shortly.⁸ The second sense is that which is *not* open to doubt and does not allow for being mistaken. In this case, the question arises whether it is appropriate to use "I know" at all.

² G. E. Moore, "A Defence of Common Sense," in *Philosophical Papers* (London: Allen & Unwin, 1959), 33-34

³ *Ibid.*, 38-42.

⁴ *Ibid.*, 40.

⁵ *Ibid.*, 42-43. Because again an assumption is made on the part of the sceptic, of there being knowledge at one point or another of something they are denying to exist; and this is very similar to Wittgenstein.

⁶ Though of course not entirely novel, his way of expressing it, in terms of his common-sense-arguments, is enlightening.

⁷ Norman Malcolm indeed notes this in his 1949 criticism of Moore, "Defending Common Sense", in which he states, "I believe that [...] Moore misused the expression 'I know' [...] I wish to show Moore's use [...] is contrary to [the] ordinary and correct use." See Norman Malcolm, "Defending Common Sense," *The Philosophical Review*, Vol. 58, No. 3 (1949), 202.

⁸ Akin to Plato's conception of 'Justified True Belief' (JTB).

On Certainty picks up as a response, particularly to Moore's infamous proof "Here is a hand."⁹ Wittgenstein believes Moore goes wrong in a fundamental way in his use of "I know" The first sense of 'know' is that everyday sense in which we use "I know" as "*I am certain.*"¹⁰ As Wittgenstein discusses in §18:

"I know" often means: I have the proper grounds for my statement. So if the other person is acquainted with the language game, he would admit that I know. The other, if he is acquainted with the language-game, must be able to imagine how, one may know something of this kind.

In this case, knowing is susceptible to an objective test or evidence. This is the knowing for which it makes sense to take pause and examine the evidence supporting our proposition. This is to say that it is *evidentiary*.

Indeed, as Wittgenstein states, "Whether a proposition can turn out false after all depends on what I make count as determinants for that proposition."¹¹ By "determinants", Wittgenstein means the factors by which we define truth and falsity. To illustrate, let us take Moore's proposition—"The earth existed for a long time before my birth"¹²—and line it up beside another proposition—"This tyre here is damaged". What do we take as determinants for these propositions? In the first case there are no determinants; it is simply taken for granted. It is only in the second case that we can imagine bending down and directing our friend to that spot at which a nail has punctured the tyre. Upon examination, our friend can either say "yes, here is the nail, you are right", or "there is no flaw, you are mistaken." We understand what it would look like to be mistaken in this case, realizing that we mistook a mark for a puncture upon closer examination. And this admits of some degree of certainty in this sense of 'know', for we realize we were perhaps only fairly certain in stating that the tyre was damaged. It is here that the nuanced first case of 'know', as justified true belief, comes out. While justified true belief is open to doubt, susceptible to evidence, and is open to being mistaken,

⁹ G. E. Moore, *Proof of An External World* (London: H. Milford, 1939).

¹⁰ Ludwig Wittgenstein, *On Certainty*, eds. G. E. M. Anscombe and G. H. von Wright (New York: Harper & Row, 1972), §8. Emphasis added.

¹¹ *Ibid.*, §5.

¹² *Ibid.*, §469.

it does not admit of degree because it is true—some justification has been found for its truth. It is a peculiarity of knowledge, if one does actually *know* it, that it has to be true.¹³

With Moore's proposition however, it is not clear how we could even possibly be wrong, since the doubt itself seems to be empty.¹⁴ As Wittgenstein notes in §4, “But what about such a proposition as ‘I know I have a brain’? Can I doubt it? Grounds for doubt are lacking! Everything speaks in its favour, nothing against it. Nevertheless it is imaginable that my skull should turn out empty when it was operated on.” Wittgenstein is telling us that just because we can imagine an empty skull, it does not mean that we actually believe in the possibility that we have an empty skull, and consequently, that we can meaningfully doubt that we have a brain. If we cannot meaningfully doubt, it is also meaningless to say “I know”. This is bolstered in the simple statement of §14: “That he does know takes some shewing.” As Deborah Orr notes: “Moore [...] does not ‘know’ in the ordinary sense of having *substantiated* his beliefs with some sort of justification.”¹⁵ If we are to claim knowledge, it has to be in the first sense of “know,” where it is open to evidence.

At this point, the second sense of ‘know’ emerges clearly as one in which the grounds for doubt are lacking. If a man with bandaged hands tells us that he “knows” his hands are under the bandage, we believe the man’s verification only in the sense in which we understand how he could have checked that his hands were still attached to his arms—by removing the bandages from them.¹⁶ In this case it makes sense for us to have a *localized* doubt of the existence of his hands—a doubt that allows for some sort of verification. But for Moore’s sceptic, who doubts whether those hands (and other physical objects) exist at all, the same type of verification is not possible. Wittgenstein wants us to avoid Moore’s temptation of replying to the sceptic’s charge here with the claim “I know...”, for this sense of ‘know’ is non-evidentiary, for “From it seeming to me—or to everyone—to be so, it doesn’t follow that it is so. What we can ask is whether it can make sense to doubt it.”¹⁷ We are no more certain that a hand exists in the world after presenting it than we

¹³ *Ibid.*, 415.

¹⁴ “What we call historical evidence points to the existence of the earth a long time before my birth;--the opposite hypothesis has *nothing* on its side.” *Ibid.*, §190.

¹⁵ Deborah Orr, “Did Wittgenstein Have a Theory of Hinge Propositions?,” *Philosophical Investigations*, Vol. 12, No. 2 (1989), 142.

¹⁶ Wittgenstein, *On Certainty*, §23.

¹⁷ *Ibid.*, §2.

were before. Aside from the idea of “look closer”¹⁸, we cannot imagine what it would be like to admit of mistake with the proposition “I know here is a hand”. Thus, with propositions of this kind, we are not open to being mistaken (unless there is some misunderstanding in the foundation of our language-game, or one is mentally deranged).¹⁹ If we cannot be mistaken, it becomes questionable whether it is sensical to say we can “know” at all. Moore wrongly includes, on Wittgenstein’s account, the examination of his own hand in the category of “knowledge”, since it is not open to evidentiary examination. Therefore, it does not accomplish what Moore intends in responding to the sceptic, but simply raises the further question of *how* he knows what he claims to know. Wittgenstein thinks this is a pointless pursuit and wants to go in an entirely different direction by leaving ‘hinges’ to support our empirical propositions while themselves being excluded in the realm of knowledge.

For Wittgenstein, where the sceptic ultimately goes wrong is in assuming that these types of propositions—hinges—are open to doubt. It is a mistake to think “Here is a hand” is an empirical proposition. Rather, we must understand that it plays the role of a logical proposition. It is in questioning the “river-bed” on which our language-games are built that the sceptic undermines the very rational framework, the common ground of discourse, within which doubt is postulated to begin with.²⁰

I find Wittgenstein’s parsing of the two concepts of ‘know’ to be enlightening in regard to the way we understand sceptical challenges in philosophy. It certainly identifies a misstep by Moore that we must avoid if we are to respond sensically in the face of such challenges. In breaking these two concepts of knowing apart, he highlights the unison of knowledge and doubt—of import epistemologically. In determining what counts for knowledge, it is helpful to follow Wittgenstein in only allowing that description in cases where there can be sensical doubt, with the possibility of evidentiary examination, and the potential for either truth or falsity. To me, it makes a great deal of sense that doubt is a local practice, and that with the introduction of ‘hinge’ or ‘river-bed’ propositions, there are things beyond sensical doubt.²¹ Perhaps one of

¹⁸ *Ibid.*, §349.

¹⁹ Here we can understand “language-games” in the same sense Wittgenstein used them in the *Philosophical Investigations*, namely, as our basic linguistic interactions, along with the rules structuring those interactions, and the actions accompanying them.

²⁰ *Ibid.*, §§97, 99, 341.

²¹ *Ibid.*

the best examples of this is the child learning a language game, summed up nicely in Wittgenstein's rhetorical questions in §478: "Does a child believe that milk exists? Or does it know that milk exists? Does a cat know that a mouse exists?" I agree with Orr that Wittgenstein "does not hold that having beliefs necessarily involves propositions", rather there is a "pre-linguistic" element, in which "belief...is shown by the behaviour and reactions of those beings involved and is entirely unconnected with linguistic achievement."²² If we treat "milk exists" as a hinge, I think the child's conviction cannot be justified; it is something that is animalistic. Moyal-Sharrock would share this view in suggesting that there are basic animal "certainties" we possess that cannot be known, but can only be shown in acting.²³ I think both Orr and Moyal-Sharrock are right in arguing that some possession of a "picture" which underlies the belief, or system of basic beliefs, by the aforementioned child, is mistaken.²⁴ Wittgenstein is here "discussing a kind of certainty which is part of the human form of life and which relates to non-doubting behaviour."²⁵ So "milk exists" is "non-propositional...it relates to non-linguistic human...experience."²⁶ We cannot ignore, as Orr notes, the crucial role of "pre-linguistic" and "non-linguistic" experience in *On Certainty*.²⁷

In summation, Wittgenstein is saying that propositions such as Moore's should not be rightfully counted as part of our knowledge. As Malcolm states, "They are not propositions that are 'known' or 'not-known'. Their peculiar role is to define the boundaries within which we raise questions, make investigations, conjecture, verify, reason. Without such boundaries we would be speechless and thoughtless."²⁸

Given the above discussion, Wittgenstein thinks it mistaken to treat 'hinges' with the language of truth and falsity. It is the very point about 'hinges' that "whenever we test anything, we are already presupposing something that is not tested."²⁹ Having "learnt" something reminds us of the "groundlessness of our believing."³⁰ This "groundlessness" is the complex system of hinges which underlie our

²² Orr, "Did Wittgenstein Have a Theory of Hinge Propositions?," 140.

²³ Daniele Moyal-Sharrock, *Readings of Wittgenstein's On Certainty* (Hounds-mills, Basingstoke, Hampshire: Palgrave Macmillan, 2005), 94.

²⁴ Orr, "Did Wittgenstein Have a Theory of Hinge Propositions?," 140.

²⁵ *Ibid.* The passage Orr is specifically referring to in *On Certainty* is §354.

²⁶ *Ibid.*, 141.

²⁷ *Ibid.*

²⁸ Norman Malcolm, "Moore and Wittgenstein on the Sense of 'I know'," in *Thought and Knowledge: Essays* (Ithaca: Cornell University Press, 1977), 197.

²⁹ Wittgenstein, *On Certainty*, §163.

³⁰ *Ibid.*, §§165-166.

language games but eschew propositional assignment. This system, our “world-picture,”³¹ which is inherited, is made up of ‘hinges,’ which, as Coliva states: “simply are not in the business of semantic evaluation.”³² Wittgenstein’s point is that our system of verification is contained within our language game. Therefore, any discussion of determining truth necessitates our prior acceptance that we are dealing with statements that can be propositional, namely, *not* hinges.³³

Now, there are exceptions to this over time, as contexts change. Indeed we have been provided with an example since Wittgenstein’s death, namely, that the proposition “Nobody has ever been on the Moon!” is now false. But it would be mistaken here to think that this means these hinges are, in some cases, open to such an evaluation; rather, in such cases these sentences cease to act as ‘hinges’ for us in our language game. As Coliva states, “the fact that we no longer except ‘Nobody has ever been on the Moon’ from semantic evaluation...simply shows that that sentence no longer plays a normative role for us.”³⁴ Yet, the thought mentioned at the outset of this paper—that we cannot say we *know* that ‘the Earth has existed for a long time before my birth’—seems puzzling. The distinction in approaching this issue is that it is not a matter of whether the collections of history, the advances of the geological sciences, et al., can bring such a sentence into the light of evidentiary examination, and therefore render it a candidate for knowledge. Rather, this sentence is in that category of ‘knowing’, independent of doubting, which is not ‘knowing’ at all; it is a *taking for granted* that is incompatible with truth values. This is to pick up on a grammatical use of “know” which stands in contrast to the epistemic use.³⁵ The former is hollow insofar as it is an expression of our certainty of something which cannot be doubted; therefore, it is arguably a mistaken use of “know” in the true epistemic use of knowing-something or knowing-that, in which there is a justificatory relationship between subject and proposition.

I will address at this juncture the issue of continuing to refer to ‘hinges’ as propositions. I think it is a simple happenstance of our

³¹ *Ibid.*, §93.

³² Annalisa Coliva, *Moore and Wittgenstein: Scepticism, Certainty, and Common Sense* (Hounds mills, Basingstoke, Hampshire: Palgrave Macmillan, 2010), 183.

³³ Wittgenstein, *On Certainty*, §§197-198.

³⁴ Coliva, *Moore and Wittgenstein*, 184.

³⁵ Wittgenstein, *On Certainty*, §1-6.

vocabulary that we refer to hinges as “propositions”, and this should not be seen as having some deeper meaning. To suggest a reading of Wittgenstein that allows ‘hinges’ as propositional creates a conflict with the Wittgenstein of the *Philosophical Investigations*. Howard Mounce argues, “‘Hinge propositions’ are readily understood once they are seen not as propositions existing unstated in a practice but as what a philosopher would say who wished to make explicit its essential features.”³⁶ I see this as misguided because it allows “philosophical” uses which are nevertheless propositional. ‘Hinges’ are instead normative in nature, making them meta-linguistic, but still a part of our language and not outside of it. As Moyal-Sharrock says, they are “certainties that can only show themselves in what we say and do.”³⁷ This is to stress that they are not some system of basic beliefs, which underlie our language, but rather a collection of animal certainties that are the basic psychological underpinnings of linguistic interactions. Thus, just as Wittgenstein suggests doing away with our confused sense of “I know...,”³⁸ I think it as fitting to refer to these *hinge-propositions* as simply ‘hinges’, as I have already begun to do in this paper.³⁹

So, since our system of verification is contained within our language game such that the discussed ‘hinges’ or ‘bedrock’ propositions themselves determine the bounds of truth and falsity, these propositions therefore cannot then be subjected to judgments of truth and falsity. Yet, at *OC* §83, Wittgenstein states “The *truth* of certain empirical propositions belongs to our frame of reference.” And again, at §206, “If someone asked us ‘but is that true?’ we might say ‘yes’ to him; and if he demanded grounds we might say ‘I can’t give you any grounds, but if you learn more you too will think the same.’” Now, in these paragraphs, Wittgenstein neither affirms ‘hinges’ as true, nor considers them false; instead, as Coliva suggests, he tells us that they indicate a “rule” such that “the truth of it [hinges] is incontestable [for us].”⁴⁰ However, this is where I think the nuance of justified true belief is interesting, and perhaps Coliva does not push far enough. For, in *OC* §505, Wittgenstein states, “It is always by favour of Nature that one knows something.” And we will remember the “grammatical peculiarity that ‘p’ follows from ‘I know p’.”⁴¹ Within a language-game using “know” in the grammatical

³⁶ Moyal-Sharrock, *Readings of Wittgenstein's On Certainty*, 118.

³⁷ *Ibid.*, 94.

³⁸ Wittgenstein, *On Certainty*, §415 and following.

³⁹ And I mostly do so throughout.

⁴⁰ Coliva, *Readings of Wittgenstein's On Certainty*, 183.

⁴¹ Wittgenstein, *On Certainty*, §415.

non-epistemic sense can introduce a ‘hinge’. In doing so we are simply identifying something which cannot be mistaken, for no system of verification exists beyond our certainty of it. But I want to argue that it is important to observe the distinction Wittgenstein suggests between *being* true and *showing* something is true. Being true *within* a language game is different from *being* true.

Coliva seems to suggest that truth ends within the parameters of our language game, but the above observation shows that this is insufficient. Rather, Wittgenstein ultimately seems to place truth beyond our language game, such that Nature has the final say on matters of truth; and this will be uncovered not in what we see, but in *acting*.^{42,43} This is to read Wittgenstein as neither a classical realist nor an idealist, but to suggest that truth relates to the certainties which make up ‘hinges’ in the same way knowledge does; namely, there is a verifiable version of truth within our language games, but there is a non-evidentiary realm of truth that cannot be discussed sensically in relation to our language-games. This is the type of “truth” which we can only discover through making mistakes in the world, like stepping off a high ledge as a child, ignorant of gravity, only to decide never to do that again. Therefore, while we cannot assess ‘hinges’ to be true (or false), much in the same way it is nonsensical to consider them part of our knowledge, I believe it is more meaningful to think of their relationship to truth and falsity in terms of a current that runs through Wittgenstein’s work: a prescribed *silence*. Just as it is meaningless to ascribe truth or falsity to ‘hinges’, it is meaningless to state such ‘hinges’ at all outside of a natural context; therefore, we should not engage in the nonsensical practice of stating these ‘hinges’ in isolation, as if there was some special philosophical meaning; we simply must remain silent.

In conclusion, ‘hinges’, amongst their other names,⁴⁴ make up a family of concepts in Wittgenstein’s *On Certainty* that are elemental in his final picture of language. They are illustrative of his notion that our language is not predicated on any necessary knowledge, and that there are certain pre-linguistic and non-linguistic commitments that we make in order to learn and function within our language games. I have argued that, when it comes to discussing ‘hinges’, instead of finding the terminus of truth and falsity within our language games, we must

⁴² *Ibid.*, §204.

⁴³ I think Nature is the arbiter of truth in this case.

⁴⁴ *Ibid.*, §§152, 211, 401-403, 141-142, 95.

understand that truth lies beyond. This is not to conjure a meta-language; rather, it is to make the necessary distinction between *showing* something to be true, and something *being* true. Therefore, I have forwarded a reading based on the Wittgensteinian prescription of silence on certain matters, as one that is fruitful in the case of examining the relationship between ‘hinges’ and truth and falsity.

Works Cited

Coliva, Annalisa. *Moore and Wittgenstein: Scepticism, Certainty, and Common Sense*. Hounds Mills, Basingstoke, Hampshire: Palgrave Macmillan, 2010.

Cook, John W. "Notes on Wittgenstein's *On Certainty*." *Philosophical Investigations*, Vol. 3, No. 4 (1980): 15-37.

Cook, John W. "Metaphysics of Wittgenstein's *On Certainty*." *Philosophical Investigations*, Vol. 8, No. 2 (1985): 81-119.

Hacker, P. M. S. *Wittgenstein's Place in Twentieth Century Analytic Philosophy*. Cambridge: Blackwell, 1996.

Malcolm, Norman. "Moore and Wittgenstein on the Sense of 'I know'." In *Thought and Knowledge: Essays*. Ithaca: Cornell University Press, Ithaca, 1977.

Malcolm, Norman. "Defending Common Sense." *The Philosophical Review*, Vol. 58, No. 3 (1949): 201-220.

Moore, G. E. *Proof of An External World*. London: H. Milford, 1939.

Moore, G. E. "A defence of Common Sense." In *Philosophical Papers*. London: Allen & Unwin, 1959.

Moyal-Sharrock, Daniele. *Readings of Wittgenstein's On Certainty*. Hounds Mills, Basingstoke, Hampshire: Palgrave Macmillan, 2005.

Orr, Deborah. "Did Wittgenstein Have a Theory of Hinge Propositions?." *Philosophical Investigations*, Vol. 12, No. 2 (1989): 134-153.

Wittgenstein, Ludwig. *On Certainty*, edited by G. E. M. Anscombe and G. H. von Wright. New York: Harper & Row, 1972.

A Defense of Abortion Rights on Lockean Principles

Edward Ruse

Although Locke explicitly condemns the procurement of *any* abortion as immoral, I argue that his theories of morality and of personal identity implicitly permit exceptions to this ban. I propose that Locke's absolute condemnation of abortion results from the context in which he is writing and from an appeal to Revelation, which opposes the greater themes held throughout his works *An Essay Concerning Human Understanding* and *Two Treatises of Government* (hereafter referred to as the *Essay* and the *Two Treatises*). Hence, we might interpret Locke's absolute ban on abortion as a pragmatic clause to avoid morally impermissible abortions, rather than a conviction held strictly on the basis of principles. I argue that Locke did not consider the possibilities that would be made available by advances in medical technology and that if he had, he would have permitted therapeutic abortion in two types of cases: those in which the mother's life is at great risk, and those in which the fetus is so damaged or underdeveloped that it cannot develop into a person. These exceptions follow from his principles of Natural Law and his criteria for personhood, respectively.¹

I. Context: Religious Time of Poor Technology

To begin, it is important to understand that the context in which Locke wrote is inseparable from the convictions he holds throughout the *Essay* and the *Two Treatises*. Locke wrote in a biblical framework and in a time with medical technology far inferior to that in the world we live in today; his ban on abortion is a function of these two factors.

First, nearly every normative statement Locke makes has its foundation in Revelation and biblical text. From the duties we have as persons to abstain from committing suicide, to our inability to justly

¹ The concept of 'person' to Locke will be addressed in the section, 'Abortion of Non-Potential Persons'. At present, it is sufficient to understand that a person is a rights-bearer.

harm the well-being of the innocent, to our duty to provide charity to others, all are in some sense derived from the idea that we are an intelligent being's creation and rightly His property; a ban on abortion comes from this same rationale. From the bible: "Now the word of the Lord came to me saying, 'Before I formed you in the womb I knew you, and before you were born I consecrated you.'"² Sections like this have historically been cited to grant legitimacy to the idea that a fetus holds the same right-bearing status as adult. When viewed in this framework, banning abortion seems a just and rational course of action. If a fetus is considered to be God's property, then it logically follows that it should have similar protection from murder as any adult, as all are God's property and protected by the Natural Laws—this is how abortion is viewed by many religions. We have a duty not to kill innocent persons in Locke's model, and for all intents and purposes fetuses have historically been categorized as innocent persons or potential-persons by the religious community. It follows, given what we know about Locke's life, that his ideas would reflect and sympathize with such a view.³ However, it does not follow soundly from the principles which hold Locke's system together that such a ban should be absolute.

The second major factor to be taken into consideration regarding the context of Locke's writing is the state of medical technology during his time. Whereas today such devices as ultrasound machines and medical monitoring equipment grant the ability to accurately measure the health both of fetuses in the womb and of pregnant mothers, such technology in Locke's time was no more than a work of fiction. While intuitively it does not seem as though technological advances should affect the principles of Locke's theories, I believe that the lack of such devices led Locke to take a pragmatic stance on abortion and to ban the practice in order to avoid abortions of healthy fetuses and abortions procured on mothers whose lives were only seemingly at risk.

It may be thought, given Locke's commitment to medical research during his life,⁴ that he would have considered the possibility of the development of better technology to monitor the health of a mother

² Jer. 1:4-5 New Revised Standard Version.

³ Locke lived in a highly Christian time in England's history and was raised by Puritan parents. Furthermore, he went to Oxford for his formal education (at the time was a religious institution—causing Locke to even consider the life of a Clergyman). William Uzgalis, "John Locke," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, §1.1.

⁴ Uzgalis, "John Locke."

A DEFENSE OF ABORTION RIGHTS...

and fetus during pregnancy. However, given how remote even the discoveries related to electricity's use were when the *Essay* was written, it is unlikely that he would have believed that the precision of modern medical devices was ever achievable.

When we consider these two aspects of the context of Locke's writing—i.e. Revelation saying that certain cases of abortion are certainly immoral, and the inexistence of technology capable of helping us determine which cases were of this kind—it appears plausible that Locke's absolute ban on abortion was a pragmatic solution to a problem that has since been solved. Due to the capabilities made possible by modern medical technology, doctors can more definitively determine the health of pregnant mothers and fetuses in the womb. I hypothesize that if such technology was available in Locke's time, then his philosophy would not be so quick to place a complete ban on procuring abortion. Instead, it appears that the ban would be limited to abortion of healthy fetuses in mothers of good health as determined by a modern medical assessment.

II. Locke's Ban on abortion

With the framework in place, a good starting point is the section in the *Essay* in which Locke explicitly bans procuring abortion. In passage 1.3.19 of the *Essay*, Locke makes a normative statement that abortion is immoral. He states, "...it is part of the Worship of God, not to kill another man; not to know more women than one; not to *procure abortion...*" (*Essay*, 1.3.19). Here he sets out a series of actions which, like the foundations for much of his theory, Revelation tells us are immoral.

While these assertions above do follow directly from Revelation, they do not seem to follow as consistently from reason in a Lockean model, and further, many of them seem to have exceptions to their bans throughout Locke's writing. For example, his statement that we are not to know more than one woman is consistent with Revelation, but opposed by reason in his assertion that we should have a right to divorce.⁵ His statement that we are not to kill another man, while supported by Revelation, is opposed by reason in his chapter, 'States of War' because of the threat that allowing a criminal to live can pose to our self-preservation.⁶

It seems consistent that abortion here is no different from the

⁵ John Locke, *Two Treatises of Government: A Critical Edition* (Cambridge: University Press), Book 2, Chap. 7, Para. 82.

⁶ *Ibid.*, Chap. 3, Para. 7.

other non-universal bans that Locke lays out as absolute. While Revelation strictly bans the act of procuring abortion, reason tells us that there exist situations in which it is permissible to infringe this rule, if the conditions are right.

III. Types of Permissible Abortion

There are two cases of abortion that I argue are implicitly permitted by principles held in the *Essay* and the *Two Treatises*. The first type of case is that in which not aborting threatens the mother's life. The second type of case is that in which evidence indicates that the fetus cannot develop into a person—that is, into a rights-bearer. Demonstrating the permissibility of abortion in this first case—which should be an uncontroversial—will serve to discredit an absolute Lockean ban on abortion: If an absolute ban on abortion follows from Locke's principles, we should find no exceptions to the ban. Given that we can, uncontroversially, find at least one exception to the ban, an absolute ban that follows from Locke's principles is unlikely.

The second type of case—the non-person case—is more controversial than the first, but given that an absolute ban on abortion is made unlikely given that we have accepted the permissibility of abortion in the first type of case, the permissibility of abortion in the second case is still rooted in a plausible and charitable interpretation of Locke.

IV. Abortion to Save the Life of the Mother: An Appeal to Natural Laws

In situations where evidence demonstrates that not aborting threatens the mother's life, Lockean Natural Law—specifically the duty to preserve one's life—not only permits, but requires procurement of abortion. Abortion in these instances fall under the category of means *to preserve oneself*: means we are obligated to pursue as each individual is “bound to preserve himself [or herself].”⁷ Saving an unborn child's life in these circumstances should be thought of as falling under our obligations to preserve the rest of mankind, an obligation which is explicitly superseded by our duty to preserve ourselves.⁸

Locke probably did not address such cases because determining the risk to the mother's life was not a developed science during his time; in modern times it is. In Locke's time, devices such as ultrasound

⁷ *Ibid.*, Chap. 2, Para. 6.

⁸ *Ibid.*

A DEFENSE OF ABORTION RIGHTS...

machines were either yet to be conceived or a work of total fiction. There were no practical means of determining whether sustaining pregnancy would endanger the mother's life. I suggest that Locke remedied this with a pragmatic solution to this problem: ban all cases of abortion. This practical solution, I hypothesize, is a reflection of the resources available to Locke for law making, rather than a conclusion rationally derived from the Natural Laws.

One might object: Locke does permit the forfeiture of one's life where "...some nobler use, than its bare Preservation calls for it."⁹ While this clause may permit a mother to sacrifice her life in certain situations in which the fetus is developed enough to survive and develop outside of the womb, this could not answer situations in which the abortion must be performed at an early stage in the fetus's development, where it could not survive independently of the mother. Even in cases of later abortion of this type, it does not seem as though Locke's philosophy adequately provides the means to precisely determine whether a mother sacrificing her life for her fetus is the noblest use of her life – perhaps she is a mother of many children and intends to bring more into this world if she survives the pregnancy. Regardless, Locke's system does not provide objective criteria for when a life may be sacrificed for some nobler use, and we can easily envision a scenario in which no nobler use can be conceived and thus that the mother must remain obliged to procure abortion to preserve her life. If we can envision such a scenario, then the apparently absolute ban on procuring abortion has been shown to have at least this exception, given Locke's Natural Laws.

V. Abortion of Non-Potential Persons: An Appeal to Locke's Notion of a Person

The second type of morally permissible abortion is abortion done when the fetus is either so highly damaged or so underdeveloped in the womb that the fetus is incapable of developing into a person, and therefore, cannot be a potential person. As we will see below, given Locke's criteria for personhood, it is reasonable to attribute to Locke the view that some fetuses are not potential-persons, and are, consequentially, not protected by the claims that follow from being a rights-bearer.

Locke views the growth of a being from conception to death as sharing the same identity throughout life. He states, "an oak growing from a plant to a great tree, and then lopped, is still the same oak."¹⁰ This

⁹ *Ibid.*

¹⁰ John Locke, *An Essay Concerning Human Understanding*, ed. Roger Woolhouse (London: Penguin Books, 1997), Book 2, Chap. 27, Para. 3.

entails that if a fetus is to grow into a person, then it too shares the claim to life that this person will possess—at this stage it is a potential person. It follows that if a fetus is to develop normally into a person, into a rights bearer, then the mother clearly does not have the privilege to decide the fetus's fate arbitrarily, as we all have a duty in regard to innocent rights-bearers not to "...harm another in Life, Health, Liberty, and Possessions."¹¹ The question, as to whether abortion in these situations is permissible, hinges then on the correct interpretation of Locke's criteria to be a person and whether certain damaged fetuses may ever meet this criteria.

Locke states that the title of 'person' is not granted for the physical constitution a being possesses, but rather for one's mental capacity to exhibit rational consciousness. He states, "...to find wherein personal identity consists, we must consider what person stands for; — which, I think, is a thinking intelligent being, that has reason and reflection, and can consider itself as itself."¹² We see by this definition of personhood that an individual must have certain cognitive faculties in order to be considered a person and consequently granted the refuge given to rights-bearers. It is also made clear by use of conjunctions in Locke's definition that to be considered a person an individual must have all the traits: being an intelligent being, having reason and reflection, and the ability to consider itself as itself; it is not sufficient to merely be in the sub-category of animals, Human, which usually possess these traits.

Locke also states: "This may show us wherein personal identity consists: not in the identity of substance, but, as I have said, in the identity of consciousness."¹³ We see again from this quote that merely possessing a human body is not sufficient to establish one as a person, as having the physical composition of a man is never stated as a criterion for being classified as a person. Locke even goes so far as to humor the idea of a rational parrot which, while not a man, does possess the qualities of a person and seems as though would qualify as a rights-bearer.¹⁴ Clearly, Locke's criterion for personhood lies in the possession of a certain type of rational consciousness.¹⁵ This shows us that cases of

¹¹ Locke, *Two Treatises*, Book 2, Chap. 2, Para. 7.

¹² Locke, *Essay*, Book 2, Chap. 27, Para. 9.

¹³ *Ibid.*, Para. 19.

¹⁴ *Ibid.*, Para. 9.

¹⁵ It is worth noting in a footnote that Locke does consider the possibility of the existence of an immaterial soul, but is ultimately agnostic regarding its existence. If there did exist an immaterial soul within each person then clearly

A DEFENSE OF ABORTION RIGHTS...

permissible abortion of damaged fetuses must be limited to those in which the cognitive faculties are known to be incapable of developing to a sufficient level to be considered a rational and intelligent being, not cases in which the fetus is known to possess physical deformities – unless the physical deformities are such that they highly affect the mental capacity of the child-to-be.

The line dividing those who can be considered persons, and those who cannot is inherently hard to draw. Precisely how much consciousness is required for this individual to be considered a person is unclear. Nevertheless, we can say that there must exist a baseline mental capacity that a human must reach to even be considered a person, given Locke's criteria. To illustrate my point, I propose a generously low threshold of consciousness—some function of minimal awareness and capacity for self-reflection—that must be met to be a person: if a fetus can be seen to possess qualities that indicates it will eventually meet this threshold, then it is a potential-person.

There exist situations in which a fetus is so damaged or underdeveloped that it is incapable of ever developing a level of brain activity sufficient to meet the threshold for personhood. In these situations the fetus is not a potential-person, as it is known not to have the capacity to develop in the future into a person. For Locke's time, it is clear that the pragmatic answer to such circumstances would be to prohibit the procurement of abortion because of the inability to definitively know if a fetus is a potential person or not. With current medical technology this qualm has since been resolved.

If a fetus is not a potential person, it is clearly not protected by the claims that belong to normal rights-bearers. The question turns then to the relationship the mother has to the immutably damaged fetus in such scenarios.

VI. A Damaged Fetus as the Property of the Mother, Vital Union Argument

personal identity would not lie in the possession of a brain capable of rational consciousness, but rather in the virtue of merely being a part of the human race. However, being that Locke is undecided on the ultimate existence of a soul, he has constructed a system in which something more akin to adequate brain function and the ability to exhibit rational consciousness are the distinguishing features of a person so to avoid placing a foundation of personhood in something he could not prove to exist. Because of this, the principles of his philosophy should be judged with respect to personhood consisting in the possession rational consciousness, not in the existence of an immaterial soul.

After a fetus has been determined to not be a potential person, providing the means to justifying its abortion is a simpler task. The fetus here has shifted from a potential person to a mere substance. At this point, the fetus can be shown to be the definitive property of the mother with justification of its disposal being analogous to the justification behind an amputation of an appendage or any other non-essential part of an individual's body.

Amputations and other non-fatal surgeries which remove parts of a person's body (think appendicitis surgeries) are justified in the Lockean model as a disposal of one's property. For something to be considered property in this respect, it must share a special relationship with the owner's body as illustrated in the following quote: "...wherein the identity of the same man [or woman] consists; viz. in nothing but a participation of the same continued life, by constantly fleeting particles of matter, in succession vitally united to the same organized body."¹⁶ This quote tells us that by Locke's account, if something is "vitally united" to a person then it is part of her identity—consequently, it is her property to dispose of as she pleases so long as its disposal would not result in the forfeiture of her life.¹⁷ This concept that, if something is connected to a person in such a way as to be part of their person it is then considered their property, is reaffirmed directly in *Two Treatises* when Locke states that "...every Man has a *Property* in his own *Person*."¹⁸

It follows by this account, if a damaged fetus is vitally united to the mother who houses it, then it is her property to dispose of as she pleases. It seems as though a fetus is vitally united to the mother in this respect as it is the mother's body that sustains and repairs the fetus as if it were any other appendage or extension of her body. There is a constant exchange of fleeting particles of matter between the two which seem to function as though they are part of the same body. The mother's body provides the fetus with oxygen, nutrients, and antibodies while simultaneously regulating growth by circulating hormones throughout the fetus. Reciprocally, the fetus has actually been discovered to produce hormones that physiologically affect the mother.¹⁹ The relationship here is analogous to that between a person and any major organ of the body, it

¹⁶ *Ibid.*, Para. 6.

¹⁷ Locke, *Second Treatise*, §2.6.

¹⁸ *Ibid.*, §5.27.

¹⁹ Gerard M. Dileo, "Anatomy of a Fetus: The Placenta," accessed December 1, 2013, http://www.babyzone.com/pregnancy/fetal-development/anatomy-fetus-placenta_70439.

A DEFENSE OF ABORTION RIGHTS...

follows that the relationship should be categorized as a vital union.

I believe this connection is generally overlooked by Locke scholars because of the special relationship a fetus of normal health shares with the mother who houses it. In situations of normal house, a fetus is both vitally united to the mother, yet protected by the claims entitled to any rights-bearer—specifically a claim to life, so that any claim to property of the fetus by the mother is clearly superseded by the fetus's claim to life. However, in situations in which the fetus is known to not be a potential-person, it appears as though its status is then lowered to merely being a substance which is vitally united to the mother, and hence rightly her property to dispose of as she pleases.

If the above arguments are accepted, we are left with the conclusion that certain cases of therapeutic abortion are morally permissible given Lockean moral principles and criteria for personhood. Locke does explicitly ban abortion, probably to remain committed to Revelation as a foundation for his overarching philosophy and to pragmatically protect the lives of some potential-persons. However, a holistic reading of Locke shows us that he cannot fully be committed to this stance while consistently holding the rest of his clauses as binding. In modern times, in which medical technology provides knowledge of fetus and mother health inaccessible during Locke's era, the pragmatic stance he takes is no longer justified.

Works Cited

Dileo, Gerard M. "Anatomy of a Fetus: The Placenta." Accessed December 1, 2013. http://www.babyzone.com/pregnancy/fetal-development/anatomy-fetus-placenta_70439.

Locke, John. *Two Treatises of Government: A Critical Edition* (Cambridge: University Press, 1960).

Locke, John.. *An Essay Concerning Human Understanding*, edited by Roger Woolhouse (London: Penguin Books, 1997).

Uzgalis, William. "John Locke." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Substantive revision July 10, 2012. <http://plato.stanford.edu/entries/locke/>.

Classics and Philosophy: In conversation with Brad Inwood

Rory Harder, Brad Inwood, & SuJung Lee

Introduction: Brad Inwood began teaching at U of T in 1982; he began in Classics and later on became a part of the Philosophy department as well. He is one of the leading scholars in in Ancient Greek and Roman philosophy, and has written about and taught courses on Plato, Aristotle, and Hellenistic philosophy. His depth of knowledge in these fields earned him the position of Canada Research Chair in Ancient Philosophy in the year 2000. He is currently the Acting Chair of the Philosophy department and also preparing to leave for Yale after this academic year. Students and colleagues alike know Professor Inwood as a kind, encouraging and extremely insightful scholar; he is a valuable member of the U of T philosophic community and will be dearly missed. We decided to interview him to hear some of his reflections on his time spent at U of T, and also, of course, to learn about his views on philosophy, especially regarding its history.

Noēsis: We understand that you've been teaching at U of T since 1982; that's over 30 years now! In all these years, you must've witnessed and experienced many things. Are there any particular experiences that you will cherish as you leave for Yale? Perhaps a course that you particularly enjoyed teaching, or any specific moments of inspiration from your colleagues or students?

Inwood: There are lots, of course. I started out teaching in Classics in 1982—and half my teaching is still in Classics. Although I did teach some philosophical material right at the beginning, I didn't really become closely associated with the Philosophy department until the late 80's, early 90's. So it's been quite a transition for me. Naturally, some of

my most vivid memories are from my earliest years. The first year I taught here, I taught Plato's *Republic* in Greek, for a semester, which was a wonderful way to start. I co-taught it with a senior colleague, a soon-to-retire Classics professor who'd been teaching Plato for *his* whole career; so I felt that that was a passing-the-torch kind of moment; and I've been teaching Plato, in Greek, at the senior undergraduate level ever since. That has been great. I really opened out the course and started getting philosophy students more and more as time went on, and one of the really satisfying things in the last few years has been getting philosophy majors and specialists, who've also learned Greek, into those Plato courses; that's been very special because it has its roots, as I do, in Classics and grows out into Philosophy.

Another big transition that means a huge amount to me was when I started teaching courses *in* the Philosophy department, and that was, I think, in 1991. I had always taught a little survey course in ancient philosophy in the Classics department. And at one point, I realized if I wanted to teach philosophy undergraduates, then I had to teach philosophy courses. So I asked permission to teach the philosophy course, PHL200—PHI200 when it existed—so that I could get access to the philosophy undergraduates. And that was a very big change, just because the interests, the enthusiasms, and the skills of the philosophy undergraduates were different from those of Classics students. It was a very satisfying experience as a teacher and I enjoyed making that transition. It was the point where the philosophy department basically welcomed me aboard and made me feel a part of the community here and that's been about 25 years. In a way, that's more significant than the total length of time at U of T; that ability to make a transition into Philosophy and to realize that as a department, Philosophy has always been very open, porous, interested in other units, connected all over the place—that is great.

I've also enjoyed co-teaching that course with colleagues in Philosophy—most recently with Professor Whiting—and that experience of co-teaching a course with a colleague where you've got intellectual compatibility—lots of things different, lots of things overlapping—that's really exciting. So that ability to co-teach has also meant a lot.

There are lots of things I could go on about, including teaching introductory Greek. When I was just a beginner full of enthusiasm, I set a term test one day and really botched it; I accidentally set a version of the test that was nearly twice as long as I meant to set! Those poor students had to sit through it and I wondered why they seemed to be sweating so

much; they seemed to be so disturbed. And then I looked at the test and realized that it was the long version and not the short version. Fortunately, that was a *long* time ago! But that was memorable because I've remembered ever since to double-check my test scripts. The students all passed anyway, which was *amazing*. But that was memorable, to be sure.

Noēsis: It also speaks to your teaching abilities!

Inwood: And a bit of absent-mindedness, yes...

Noēsis: You once told your students that if you had to pick one out of all of Plato's dialogues to teach to students, it would be *Meno*. I thought that this was a great choice, and that your response would be great to share with our readers. So what's so great about *Meno*?

Inwood: Hmm... lots of reasons. But the exact question I was answering was *slightly* different, because I said that I wanted to rule out multi-volume, multi-book works. So maybe if I could do the whole *Republic* (one dialogue but *ten* books), that would be the choice.

The *Meno* is about knowledge. It asks hard questions about knowledge, about the method for getting knowledge, the limitations on knowledge, and the importance of knowledge. For me, that's always been the core of what philosophy is. I sometimes think people approach philosophy in one of two (or maybe three or four) ways. One of these is a passionate interest in metaphysics and science (what is the world actually made of, et cetera). But sometimes people start with the worry about knowing things, and that's always how I've looked at it. I've always been brought to philosophy by worries and puzzles about knowledge: how we get it, what its limitations are... And the *Meno* tackles that in a really direct way and proposes answers that have been very influential. I think the idea of tying down true opinions by figuring out the cause is a profound insight which you can still use when you think through, say, issues in philosophy of science. And although I've got colleagues who disagree with this, one of the best things about the *Meno* is that it thinks about the difference between knowledge and mere true belief without explicitly invoking Forms; so it's a metaphysically lean exploration of knowledge, and that's something I think we can still learn a great deal from. And also it's got lots of discussion of definition at the beginning. It's like a little textbook about how to do a definition: what mistakes you should avoid, what things you have to look for, etc. And then there's the slave boy and geometry and recollection—it's just got so much about

knowledge from different points of view. So the *Meno* really has it all, and the fact that it's open-ended and *aporetic* suits my temperament; better than having answers.

Noēsis: And I think that it's also a great illustration of the Socratic method.

Inwood: Yes! It's Socrates to a T.

Noēsis: We understand that yesterday (Thursday March 12th), you participated in the 2015 faculty book launch for your latest book, *Ethics After Aristotle!* Congratulations! Can you give us a brief summary of the main thesis of your book? What was your motivation for writing it?

Inwood: Oh, motivation was easy; I was invited to give a series of lectures at Harvard and I had to come up with something to say! [laughing] I had had thoughts and concerns and interests in this area for a while, but I don't think I would ever have written a book about it if I hadn't had this external motivation.

But what it's about is, well, it's a history of philosophy book, not straight philosophy, but I did want to make the history say something interesting about philosophical themes. It seems to be that that's one of the best things you can do with the history of philosophy. And I came at it from a number of angles. The period it covers is from Aristotle's first student, Theophrastus, to his greatest commentator, Alexander of Aphrodisias. His school existed for that 500 year period, and I think it hit its peak with Alexander. And through that whole period, all of Aristotle's followers—all of the Aristotelians, or Peripatetics—took slightly different views about ethics. They didn't just comment on and parrot and expound what the master said, but they all thought for themselves, and they often disagreed with the master and thought that something he said wasn't quite right. They developed their own views. And this is often the way that the history of philosophy works. They also shifted the way they present the Peripatetic position, because they're debating with different people. One century they're debating with the Epicureans and they worry about the pleasure problem; another century they're debating with Stoics or Academics and they approach the problems in a different way. As an historian, I am very interested in the way philosophers respond to the different debates they find themselves in. So, one part of the book is to describe that historical arc, that historical development; and there are

high points and low points; I mean, there are some stretches where the school is really doing pretty dreary work and others where there's just brilliant insight, and that variation is part of what philosophy is about.

Not all philosophers are great men. Plato, Aristotle, Plotinus, Augustine, Descartes—no problem! But one thing that interests me is how much philosophy is done by followers. Platonists aren't like Plato; Aristotelians aren't as great as Aristotle; Cartesians don't agree with Descartes all the time. Kantians are all over the map; they're not doing just what Kant did, I suppose. You have the great men, like Wittgenstein, and then the little Wittgensteinians. And so that's really a part of what makes philosophy work. There are the great men and the followers, or those who are inspired by them, and the dynamic around that is something that interests me historically and philosophically. If you think of yourself as a follower of a certain general strategy of philosophy, you still have intellectual independence, you still have creativity. But there can be a tension. How far can an Aristotelian depart from Aristotle and still be an Aristotelian? How far can a Platonist go and still be a follower of Plato? There are people who think of themselves as broadly "Kantian" that Kant probably wouldn't have recognized as being in the same ballpark. So there's got to be some kind of latitude and some boundaries, and that's just interesting.

It's like what happens when you're strongly influenced by a great teacher and you start to absorb some of the key lessons they've taught you and you go away and do something with it that's all your own. And of course, the best teacher is the one that encourages you to do that. Well, that's what the history of philosophy is like. You have these great philosophers, these great teachers, who have huge impact. People *want* to follow them, but they also want to be themselves. And that tension is I think one of the most important things that goes on in the history of philosophy and in the lives of individual philosophers. And so this was the attempt is to use that stretch of history, which I've been working on for a long time, to take some of that philosophical dynamic and bring that out for a more general reader. Because it's a set of lectures, it's aimed at a relatively general audience. Most of what I've written in ancient philosophy is aimed at a pretty micro-audience—for hyper-specialists. This is meant to be just a little bit broader in scope. So that was the idea behind it. Hopefully, there's a lot more in it, but that was the big inspiration.

Noēsis: Well I think that's actually a great lead for this next question; this whole idea of the importance of the teacher-student dynamic in your creative and intellectual abilities, especially in the history of philosophy.

So one of my biggest sources of inspiration for pursuing philosophy was a teacher. She was my high-school philosophy teacher. Was there a similar figure of inspiration in your life? If not, what was it that inspired you to pursue a career in classics, and eventually philosophy? Specifically, what sparked your interest in ancient Greek philosophy?

Inwood: Well, my interest in ancient philosophy—I do as much Roman philosophy as Greek—so it's really ancient philosophy as a package. There was a very important individual who inspired me. It was my first year at Brock University, and I was pretty young at the time because of Brock's early admission program. I started out thinking, like everybody else I knew, that I would be a psychology major because the human mind is so interesting. Well, that lasted about three weeks into the fall semester. So I wound up rearranging my schedule and dropping a course, and so I needed to pick up a new course in the spring. At that point there were almost no low-level courses that were just one semester long, so how could I pick up a course? Fortunately, I had gotten to know a philosophy professor named John Mayer. He must have been a very junior professor then, but he seemed very mature and wise to me. He had taught a module in the summer course that was a part of this screening for the early admission program, so he knew me a little bit. I went to him as a kind of advisor and said, "Well, Professor Mayer, I'm stuck! I need a course." He said, "OK, I've got a solution for you. I'm teaching a course on the philosophy of science, and it's a half course and it starts in January. There's only one complication: it's a third year course." I was in the spring semester of my first year, and he let me into his course, which was completely transformative in two ways. He showed that he believed in me and that challenged me to live up to his expectations. It was a huge risk for him, and I think that that inspired me personally, and I hope it has had an effect on how I have treated people ever since. Also, it was philosophy of science, which he taught as a kind of localized epistemology. And the enthusiasm for approaching philosophical problems through a concern about issues in knowledge started right there. Even though I'd already read some Descartes, somehow that didn't do it for me, but the philosophy of science did.

Then, in my second year, I was doing a Classics degree and had to pick an outside breadth requirement. So I thought, "Ah, philosophy didn't go so badly. That was a lot of fun!" so I did a course in ancient philosophy, which was absolutely fabulous. Again, just a wonderful,

enthusiastic instructor. I still remember the way he explained Heraclitus and his theory of flux. He would come in at the beginning of class, he always had his coffee with him, and he illustrated the flux theory by gesturing with his coffee cup. He called it his metaphysical coffee cup. A very concrete, memorable image—I mean this was over 40 years ago and I still remember the way he did it. Twice in a row I got really gifted teachers. I was 17 or 18 at the time, and it stuck. I didn't actually do much philosophy, formally, but I read a lot and audited lots of courses. I actually haven't *taken* many philosophy courses. But somehow that didn't bar me from philosophy. There's that openness and intellectual generosity that philosophy is capable of. It really meant a lot to me, right from the beginning—the generosity. I don't like putting fences around things, if I can help it.

Noēsis: That's very interesting. So there was no major in philosophy that you did at Brock?

Inwood: Nope! I did a straight Classics specialist degree. Every year two full courses in Greek and two full courses in Latin. I did do one other philosophy course as an undergrad. It was a year-long course on Aristotle, which was, I think, why I developed my enthusiasm for Aristotle much earlier than for Plato. Then I came to U of T and did a master's degree in Classics. No philosophy courses. Eventually I did what turned out to be a really quite philosophically oriented PhD thesis. Nevertheless, the philosophy wasn't really self-taught. I was influenced by all kinds of people, though seldom in a formal way. I think I was very lucky. That early exposure and enthusiasm was great; SuJung, you had a teacher who started you off. You probably had something similar, Rory. I think almost everybody could tell some such story. But I was lucky enough to be able to follow through. There really is a lot of luck involved.

Noēsis: So, going on to our new question. We have heard it discussed by some fellow undergraduates how many courses in history of philosophy here focus on the history of western philosophy. While U of T is certainly a Western institution, it is also home to an extremely diverse student population. As someone who works in history of philosophy, do you think that the lack of emphasis on philosophies of other cultures is a problem, and something that U of T should aim to address?

Inwood: Well, if I put on my department chair hat I have to say something official, so I better take it off, and just talk as an individual.

Yes, of course we need to do more to address this need. But why are we in the situation we are in? Well, partly it's because this is a western institution. That's certainly part of it. But partly it's because of history. Philosophy, as it is conceptualized in the English speaking world, is in a long line of cultural development, back to the Greeks. Even the word philosophy is Greek and the particular way we do it owes more to its Greek foundations directly, and indirectly, than it does to anything else. And that's probably not just in philosophy, but in so many other forms of cultural activity.

Since I began in Classics, for a long time that seemed really natural to me. What could be more Eurocentric than studying Greek and Latin? That began to shift for me a little bit when I realized, the more I worked on history of philosophy and thought about how it related to contemporary philosophy, what a different world the ancient world was from ours. We like to emphasize—well, many people like to emphasize—how much continuity there is, how similar it is, but the more I've done history of philosophy the more I've realized that there are vast differences. There's been a continuous development, but there are real breakpoints. The advent of Christianity was a huge breakpoint. Later on, the Middle Ages were a great breakpoint. The Renaissance was a great breakpoint. So, although we're continuously connected to them, we're just not straightforward followers. So the past, as a novelist once said, is a foreign country. And when you go back to the ancient Greeks and Romans, there is a foreignness about it.

So I did get interested in the foreignness of ancient philosophy, and that made me think about the value of what some people call comparative philosophy. I've got two doctoral supervision experiences that involve something that you might call comparative philosophy. One was comparative between ancient and modern, but in the western tradition—I helped supervise a thesis of Spinoza and Stoicism, knowing almost nothing about Spinoza (but don't say that in print!). It turned out OK, but more recently I was involved with a thesis that was a comparison between ancient Stoicism and a particular strand of Daoism. This was a thesis supervised by my colleague Vincent Shen, who's in East Asian Studies and Philosophy whereas I'm in Classics and Philosophy. This was a very multicultural project: his student was a man from the Czech Republic, whose Chinese was fabulous and his Greek was really good. So, I collaborated on the supervision of his thesis, which was comparing Stoicism and Daoism, and we had to all work together because my Sinological colleagues didn't really know that much

about Greek philosophy, and I didn't know anything about Chinese philosophy. But I learned a huge amount because the theme of the thesis was the comparison: similarities and differences. And trying to explore that without being imperialist in either direction was methodologically very difficult, and I think everybody learned a huge amount. It convinced me that there's a really important dialogue that has to happen. What the implications of that are for curriculum and teaching, that's a separate issue. But in terms of the intellectual project of doing philosophy, we should be interested in comparison—intercultural comparison, in the same way, I think, that we're interested in history. These are ways of getting us out of our present heads through a dialogue partner either in a different culture now, or some other version of our culture then, or both. Just so that we're not talking to ourselves and only ourselves, to provide context.

What does this all mean for curriculum? I guess we normally think of Asian philosophy, or various forms of Asian philosophy, as the main way to enrich the curriculum that we offer, although I did have a student once who now teaches African philosophy. But why don't we do it more? Partly because institutions are glacial. Institutions owe a debt to their past. We tend to expand and get adventurous only when we have extra resources, and, as everybody knows, we're not in an environment, anywhere in North America, where universities have extra resources. So we've been retrenching to preserve what we can and to pay our debt to the past, and that's limiting. If we weren't, basically, being slowly starved as a university system, I think we'd have a lot more forward-looking, outward-looking elasticity. It's understandable that we've tended to retrench, but I think it's also a mistake. The other constraint is the law of supply and demand. You only want absolutely first-rate scholars teaching at a place like U of T, or really any university. I mean, U of T is great, it should be greater, but you only want first-rate thinkers teaching at universities. And there's demand for people who can teach Asian philosophy, with not much supply by comparison. So finding first-rate people is hard. I think we'd do a great deal more if there were more being trained. But then there's this kind of vicious circle: how do they get trained? You need more faculty. So, again, the answer is having enough resources, enough support, to be able to launch something new. Most institutions are conservative, and they're protective. They're not going to jettison something that they've done for 150 years just to add something new. Alas! Which branch of philosophy would be made to lose three positions so that you could get three of something else? So, I think that's why we're in the position we're in. And obviously we should expand, that should be a high priority for when we do expand, but we're

not expanding. We're hanging on by the skin of our teeth.

Noēsis: So I think that's a natural segue to our last question. So, nowadays, philosophy undergrads are constantly being warned that there are very few jobs available in academia, and that we should only pursue a graduate degree if we are willing to accept the high risk of unemployment. Do you have any words of wisdom for those who are intimidated by this kind of advice?

Inwood: Wisdom? No. Thoughts? Sure. I think unemployment is the wrong word: *Redeployment*. It's not as though people who do graduate degrees in the humanities, especially philosophy and don't go on to tenure-stream academic careers are very often unemployed. What you get from an advanced degree in philosophy, especially, is a fabulous training for so many things. The skills are so valuable and they are honed to such a high degree. So you're not unemployed, but you're not doing what you thought you were aiming at doing. You're redeployed, and I think that's a really important distinction because, especially when you think you're at the beginning of your working life, you tend to see a narrow range of possibilities. With only a little experience that's all you *can* see. So you think: "If I don't get that, I won't have anything." That's a very natural fear. So the advice component would be to just remember that there are many more possibilities than you may think.

That said, doing a doctoral degree means a big investment of your time, energy, intelligence, and passion. It's actually kind of hard to do a PhD unless you're really driven by your passion. It's a big investment, so the question is whether it's the right way to invest yourself. I mean, if you felt it absolutely wouldn't be worth spending five or six years doing graduate work if you couldn't teach philosophy for the rest of your life—if you really felt that—then you'd be in a tough place because the guarantees aren't there. Nobody in their right mind would say, "Oh yeah, the odds are pretty good. Go for it!"

But if it's worthwhile in itself, and it's going to give you a range of options, and if you tell yourself what I told myself when I was doing my PhD, which was, "Don't go into debt", then go for it. I did my PhD in the late seventies, and at that time the prospects of an academic career were also not good. They were terrible. I just said, "Fine, I'll do it year by year as long as I don't get into debt," and it turned out fine. But for lots of people it didn't, and I used to have a little bit of survivor guilt because I actually got a great career out of it and lots of the people I

knew in grad school who were just as good didn't. There's a lot of luck. So, if you really have the drive and the passion to do it, and you're not ruining your future by dragging it out for ten years or finishing with a huge mountain of debt, and you've got the passion, then who could say no? But it's a high-risk endeavour.

It's actually worth comparing it to other high-risk things that people pursue. I mean, think if you've studied drama all through high school, then did a degree in performing arts, and you wanted to be a professional actor and make your living on the stage. Or a professional musician. Or a professional hockey player. Of all the 18 year olds who set out to be in the NHL, for every thousand young men who want to be NHL stars, how many actually ever make it? For every 18 year old who says, "I want to be a novelist" how many can? Or a professional musician? I mean, most of them don't make it, and their lives usually aren't ruined. Still, the odds of fulfilling the big dream are low. I actually think that a humanities academic career is in some ways like that. It's driven by a passion, the odds of *complete* success are low, the spin-off benefits are huge, and eventually, if it's not going well, you could wind up saying, "When do I pull the plug?" Some of the best music teachers have been people who've wanted to make their living as musicians, but found themselves as music educators. The people who wanted to act for a living wind up in teaching careers at various levels. There's just so much more flexibility than we think. So, not to follow your passion because of fear seems to me too bad—if you've got the passion. But I don't any of this counts as wisdom.

Noēsis: We think it does! Well, this is a great way to end our discussion. Thank you so much for speaking with us.

Inwood: Thank you so much for your interest. I appreciate that.

Noēsis: For sure! Do you have anything that you'd like to say?

Inwood: As you said right at the beginning, I'm winding up my career teaching here this year. For me, there couldn't be a better way to end it than as serving the department of Philosophy. It is, and continues to be despite the challenges of this particular year, it's a real honour and pleasure to work with people here. The department's always been so generous to me, and if I'm giving a little bit back, well, that's not a bad note to go out on. It'll be hard to leave.

Noēsis: Well, you will be missed!

Inwood: Thanks!

Philosophy and Obligation: In conversation with Ashley Taylor

Charles Dalrymple-Fraser & Ashley Taylor

Introduction: Ashley Taylor has been instructing at the University of Toronto since 2014. In that time, she's led a number of lectures and seminars on topics such as the philosophy of feminism and the ethics of activism. She was also an invited speaker to this year's *Symposium on Love*, where she gave a talk on flirting. Students know Dr. Taylor to be an interested and passionate instructor, and *Noēsis* is grateful for the chance to interview her before she returns to the UK next year, and wishes Dr. Taylor all the best with her future pursuits.

Noēsis: I'll start with a list and end with a question. I understand that you did your Bachelor's Degree in philosophy and international human rights at Webster University in Missouri. You then moved out to the University of Edinburgh, where you completed a Masters in Philosophy with a thesis titled *Just Solidarity*. You then stayed in Edinburgh to complete a PhD in Philosophy with a dissertation titled *The Circumstances of Justice: A reformulation*. And now, here you are newly arrived as a lecturer at the University of Toronto.

Having done—in one sense or another—philosophy in three different institutions across three different countries, I'm curious whether you've found any significant differences in the climates or approaches between UofT and those other institutions.

Taylor: Well, I did also study in Thailand for a semester and in Russia for a semester. But, I guess the main difference that I see here is actually the community. So, it took me several months for me to figure out what it actually was I thought was missing at UofT. And then I realized it: you don't go out to the pub after every talk—it's such a drinking culture in

the United Kingdom, in academia in the United Kingdom, and *especially* in philosophy. After every talk, every reading group—and there are loads of them—everybody just goes out to the pub. And it doesn't matter if you're an undergrad, if you're a grad student, or you're a professor, or whether you even drink. A few professors also held open conversation times where they would just bring tea and snacks, inviting undergraduates, graduates, and other faculty out—so it wasn't just about drinking! [laughs]

There were a bunch of professors up for a Hume Studies conference when I was up there in my first year of my PhD, and my best drinking buddy that summer was a don at Oxford. And I didn't work with him, he was just someone there to hang out with. It's just: "Hey, we all want to do philosophy, and we want to get together and have a pint sometimes when we do that." And that's not to say there's something wrong with institutions that don't have that: I think that's something unique to the UK.

Noēsis: And did you find any of this in Webster?

Taylor: No. Well, there were faculty who would invite us to wine nights or dinners every now and then, and it would be that everyone in the department was invited. But not quite as much: it was more of a coffee culture there. Our undergraduate club would meet over coffee and professors came around a lot for various things. But, in the UK, I think—there's a lot less fear of professors and things like that: everybody is on a first name basis and that's just normal.

Noēsis: Do you think a barrier to that here might be explained by the size of our department? We're very big, so sometimes it seems hard to engage—

Taylor: Edinburgh is that big too. It's just about the same size.

Noēsis: So, what about undergraduate education? Do you think there are any substantial differences?

Taylor: Smaller seminars for senior-level undergraduates is the only thing I can think of that students here could benefit from. Edinburgh is about the same size as UofT—I taught undergraduate classes there that had 450 students enrolled. But by the time you hit third and fourth year,

it's almost never over twenty students per class. And that just makes everything a lot more conducive to various kinds of learning projects and activities, but also just talking with your professors. Everything is just a lot more discussion oriented. And here, there's this forty person class which works as a *class* but limits things in a way.

But there are definitely benefits here too, don't get me wrong. You write a lot more: you do so much writing! I think a student can go to Edinburgh—as an undergrad I wouldn't go to any university in the UK except for Oxford. The results of having a big university in the UK is that, rather than having more people marking more assignments, they assign one paper per semester—and then you have a final exam in the exam period—not a lot of writing.

Noēsis: So while we're talking about papers, why don't we move to your work if you don't mind. I understand that you have strong interests in social and political philosophy: you had a paper published last year on Solidarity,¹ and recently led a senior seminar on the ethics of activism. What are some things you find particularly interesting about these topics, or about the area in general?

Taylor: You might describe my interest in philosophy as focused on obligations of all sorts and all kinds—I'm interested in everything about them. Why are people obligated? Why do people feel obligated? What is an obligation? What are the possible ways that we can defend an obligation? How do they form? I mean, the list will never end and I will probably never stop being so fascinated by them. They puzzle me and I'm sort of interested in every aspect of them. So, I can't tell you about where I started my interests in obligation, but I can tell you about my interests in solidarity.

As an undergraduate, my very favourite professor was the head of the international human rights program I was a part of. And, I remember him telling us about various solidarity projects he had been a part of. And I remember thinking: this is incredible! This is how we all should be thinking about the world around us. And I remember being struck by this at some point in my undergrad. And I got to my masters, and somehow got a hold of a paper—a special issue of the *Journal of Social Philosophy* on solidarity. But they were all talking about different

¹ Ashley E. Taylor, "Solidarity: Obligations and Expressions," *The Journal of Political Philosophy*, Vol. 23, No. 1 (2014).

things. So that set out the foundation of my paper of trying to clarify things for them.

Noēsis: So how did you go about doing that?

Taylor: When people talk about solidarity, they usually talk about something extremely robust, right? I ought to do this because I'm a member of this union on strike, for example. And that sort of robust relationship does tend to generate obligations, I think. But then other people talk about solidarity when they talk about, say, letter writing in trying to get prisoners freed. And that strikes me as weird, because it doesn't seem like they're obligated to do it in the same way as other people talking about obligation and solidarity. So what I try to do is to pull it apart, and I call the different kinds "robust solidarity" and "expressionist solidarity". And then I give an account of the necessary components, and argue that whenever they obtain in a multi-directional or bi-directional manner, then you have the conditions for robust solidarity. And whenever people are just sort of motivated along those same sort of conditions, but it's uni-directional support for a certain kind of group then it's expressionist solidarity—so you know, the Unit 1 strike is going on. I'm not a member of Unit 1, but I can show solidarity to you guys in the expressionist sense, and try to support you in other ways. Do I owe you an obligation to participate in your group in other ways, probably not. But that's how I try to explain those two kinds of solidarity.

Noēsis: So it's a lot about being a member that distinguishes things there, in terms of the kind of solidarity in which you're engaging and that bi-directionality?

Taylor: Exactly. And what I try to do that's a little different than most of the other accounts is—instead of making membership identity focused, where many accounts think you have to be a member of an association in an historical way (ethnicity, job, whatever else), I make it very project focused. That is, whatever the goal of the group is. And I think that explains things well like the strike. Some of that solidarity might be your experience with the group, but it doesn't have to be experience; you can be very new and still have a robust solidarity obligation to strike.

Now, here is just a thought about this strike. This is totally me coming from a caring about feminist issues. Forgetting all the mess of

who isn't being taught, who isn't getting paid enough, what the faculty is saying to do, wanting to help the students... From a slightly more objective perspective, imagine what raising the minimum funding package for post-graduate students would do to increase the representation and diversity in academia. Now, You can afford to do a post-graduate degree if you come from a middle-class or upper-middle class family. If people had a living wage, how would that diversify academia? Especially in the humanities where the grants are substantially smaller? Forgetting all the politics, this strike excites me for that reason.

Noēsis: That's extremely interesting. Moving on to some other questions: A lot of students today are concerned about their options for after they finish their undergraduate degrees. On the one hand, we're being warned that there are very few jobs available in academia, and that one should only pursue a graduate degree if one is willing to accept the high risk of unemployment. On the other hand, the external job market itself seems highly saturated—

Taylor: So let me jump in on this here. The job market is tough. But so is a lot of other market; everything is competitive. I honestly think, if you're asking the question "what do I do next?", the answer probably isn't "what gives me job security next month?". The right question is something like "what will make my life go best for me?". And you can think about that long-term, but people change their minds constantly. If there is anything you can get out of grad school, out of an internship, out of a job you might take for a few years but not keep—that sounds like a project worth doing.

I think the problem that all of us are facing, at least until we hit a certain age, is that there's this thought that I need a plan on how to get there. But maybe that's not the way to look at it. Maybe what makes a life go better is what you'll take out of that project. And if your life might go better out of grad school, then that's what you ought to do.

Noēsis: I think a lot of it might also be tied up in the uncertainty of being an undergraduate, unsure what that grad experience is even like for us to have an idea if it will make our lives go better off. And at the same time, the people telling us not to go to grad school, or to go to grad school, or to go elsewhere, seem authoritative in that they've usually had access to those experiences and know what it's like enough to provide an informed perspective.

Taylor: I went low-stakes with a masters first. I figured, if I don't like the idea of a PhD, I can just walk away. But a place where you can do a masters can help it feel like it's lower stakes, to help you sit down and see what it is that you want. And for grad school, which is sometimes a six to eight year project, I think it's ridiculous to commit to that without seeing what it might be like first. I don't know if it's the best thing to do, but it's certainly what I did because I've had the same worry.

Noēsis: It's interesting to say you took the low stakes by heading into a masters, but even before that, you did something very interesting after your Bachelor's degree: you served two years as a volunteer to the U.S. Peace Corps in Kazakhstan. I was wondering if you could say something about your motivations to volunteer with the Peace Corps, rather than jumping right into grad school, and your experiences there.

Taylor: Well, I wanted to save orphans in Africa, and that was pretty much how I got there. I grew up with a volunteering culture, and I wanted to make up for some of the wrongdoing I saw my country participating in. And the other part genuinely was that I wanted to save orphans in Africa. And the Peace Corps genuinely were prepared to send me to Africa, but apparently everybody who'd asked to go to Kazakhstan dropped out. And, so I thought that this thing was so much bigger than I am, and I needed to be flexible. So, I went to Kazakhstan. And I loved it, but at the same time it was the single hardest thing I've ever done or will have to do. I lived with a local family for two and a half years and I spoke Russian exclusively. Once or twice a month, my family would call, and that was my English, but that was it. I had no internet, no running water, in -40°C weather. All that aside, I did English teaching and nutritional education. I do hope some of the nutrition projects may have made some lives better off, which is the real goal.

Noēsis: So all that sounds like it was really affirming, really satisfying in a lot of senses. So, why did you return to academia?

Taylor: Well, I always planned to do a masters and go to law school. But instead I did a masters, and my primary supervisor convinced me to go ahead and apply for a PhD, just in case I decided to do so. So I applied for a practice run, and got in with funding. I felt like it was an amazing opportunity I couldn't walk away from.

Would I go back to law school? I'm not sure, but I think about it: I'm about to marry a lawyer. He's a great analytic thinker, but all his arguments are appeals to authority! [laughs]. I've heard this from friends with other philosophy degrees and law degrees, that if you go into philosophy first, it can be really hard to go into law because there are appeals to authority all over the place—there are still good analytical and critical thinking skills that are used and needed to succeed in the profession, but sort of the scope of what you need to deal with is so restricted because you're working in the law, rather than seeking the good, or truth.

Noēsis: A number of undergraduates are in the position of: "Either I'll go to grad school, or I'll go to law school." For an undergraduate deciding between the two, what do you think should be part of the deciding factors there?

Taylor: The difference to me is, do you want to be doing something in the world, or do you want to be doing something to understand and learn more? Both are worthy causes, and both need each other to truly flourish. It strikes me that you get to have a lot more real-world impact as a lawyer than a philosopher. That's not always true! There are philosophers who were excellent members of think tanks.

Noēsis: If I can return to Kazakhstan for a little bit, and your decision to return to academia, in what ways do you think your volunteering with the Peace Corps shaped the way you do philosophy?

Taylor: I've been told by various people that it's true about me. Certainly, the way I think about things that are right and wrong between nation states and international political theory has always been what's motivated me. And the way I think about what's right and wrong between nation states, and more importantly between cultures, is very much impacted by my experiences there and my ability to engage with different cultures—something that was absolutely mind-blowing at times. By and large it mostly affect my thinking about political theory and philosophy, but particularly international theory. Conceptions of good/bad and right/wrong are not as simple as we like to think they are. Some of the reason I'm so interested in social philosophy might just be because I want to deny that autonomy is the standard of thinking in a

political philosophy, and I learned that by living in a culture that doesn't put autonomy first.

Noēsis: Do you think there's a sense in which something is missing from philosophy if students don't get these experiences?

Taylor: Certainly. Some people can learn this just from reading a book somewhere. But so much of our ethical discourse is tied up in a format that says: okay, here's a principle, how does it measure against our intuitions and how do we find coherence? But our intuitions are always sort of biased toward our experiences and understanding, and to experience more to move or change our intuitions to be more objective. But I think there are more or less successful ways of doing that. For my experience, that was Kazakhstan. That helped me develop my intuitions in a way closer toward that impartial standard we're looking for.

Noēsis: Thanks so much. We've covered a lot of ground but perhaps the hardest question remains: if you could have dinner with any two dead philosophers, who would they be?

Taylor: Definitely David Hume. Absolutely David Hume.

Noēsis: Twice?

Taylor: I love him so much that maybe. Perhaps a better answer is that I'd pick David Hume and whomever he would want to invite to dinner with us, so her would talk to me more. Maybe Thomas Reid because I'd love to hear Hume respond to Reid. But Hume's the only one who matters. Philosophically, he's my first great love. He's brilliant—a lot of my PhD deals with Hume. I mean, he got a million things wrong, but I still want to give him a big hug to thank him for everything he got right. He's brilliant and nice. His view of human nature is so much closer to mine than most moral and political philosophers. When I'm having a bad day, I can kick of my shoes, grab a cup of coffee and my copy of the enquiries.

Quote on back cover: Heraclitus, as recorded in Clement, *Miscellanies* (5.140.5).

“Those who are lovers of wisdom must
be inquirers into many things indeed
—Heraclitus